

Faculdade de Engenharia da Universidade do Porto



**Exploração de ambientes sonoros
imersivos no contexto multimédia:
aplicações na mistura e desenho de som.**

Eduardo Miguel Campos Magalhães

Dissertação realizada no âmbito do
Mestrado em Multimédia - Música Interativa e Design de Som

Orientador: Prof. Dr. Carlos Guedes

Co-orientador: G. Sioros

Outubro de 2012

Exploração de ambientes sonoros imersivos no contexto multimédia: aplicações na mistura e desenho de som.

Eduardo Miguel Campos Magalhães

Dissertação realizada no âmbito do
Mestrado em Multimédia - Música Interativa e Design de Som

Aprovado em provas públicas pelo Júri:

Presidente: Doutor António Fernando Vasconcelos Cunha Castro Coelho, Professor Auxiliar da Universidade do Porto

Vogal Externo: Mestre Marco Paulo Barbosa Conceição, Assistente Convidado da Escola Superior de Música, Artes e Espetáculo, especialista de reconhecido mérito na área

Orientador: Doutor Carlos Alberto Barbosa da Cunha Mendonça Guedes, Professor Coordenador da Escola Superior de Música, Artes e Espetáculo

31 de Outubro de 2012

Resumo

A presente dissertação de mestrado centra-se na exploração de ambientes sonoros imersivos enquanto ferramenta tecnológica e estética que potencie as conteúdos visuais e narrativos num contexto criativo de *sound design* e mistura áudio.

Através da análise de e publicações científicas em campos como a ecologia acústica, percepção auditiva, cognição, psico-acústica, emoção e tecnologia musical, pretendem-se estabelecer correlações entre parâmetros e conceitos destes paradigmas e a imersão sonora do sujeito nestes ambientes. A partir desta investigação, é apresentada e descrita uma aplicação elaborada pelo autor para geração e manipulação de soundscapes em tempo real nos ambientes de programação Max/MSP e KYMA.

Abstract

This dissertation focuses on the exploration of immersive soundscapes while technological tool that enhances the aesthetics of the visual content and the narrative in the context of creative sound design and audio mixing. Through the analysis of scientific publications in such fields as acoustic ecology, auditory perception, cognition, psychoacoustics, emotion and music technology, it is intended to establish correlations between these parameters and concepts and the sound immersion of the listener in these environments. From this research, it is shown and described an application developed by the author for the generation and manipulation of soundscapes in real time in the programming environments KYMA and MaxMSP.

À Madalina e à Adriana por existirem,

Agradecimentos

Gostaria de aproveitar este espaço para incluir algumas pessoas que de uma forma mais ou menos direta contribuíram para a realização deste trabalho.

Ao meu orientador, amigo e Professor Carlos Guedes por toda a inspiração, espírito visionário e apoio quase cirúrgico em muitos momentos.

Ao meu amigo e co-orientador George Sioros pela genialidade em converter ideias e conversas soltas em perspectivas lógicas e conteúdos interessantes.

Aos meus colegas e amigos de trabalho com os quais tenho o luxo de partilhar quase diariamente momentos de verdadeira paixão pelo que fazem. Sem ordem de aparição ou preferência: George Sioros, Clara Morão, Diogo Cocharro, Rui Dias, José Alberto Gomes, Filipe Lopes, Gustavo Costa, Fabien Guyon, Marius Miron, André Holpfsafel, Rui Penha.

Ao Isidro Ribeiro e Rosário Macedo e à estrutura do DEEC pela forma como aceitaram e incrementaram os cursos de Multimédia e como contribuem para que o trabalho diário seja mais funcional e sempre com a sensação de apoio.

Aos meus pais por todo o apoio durante a minha existência e em particular durante a realização deste projeto, substituindo-me muitas vezes nas funções de pai e marido.

À Madalina pela paciência e acumulação de tarefas possibilitando-me sempre tempo para as minhas ideias em sacrifício pessoal quase constante.

Índice

RESUMO	III
ABSTRACT	IV
AGRADECIMENTOS	VI
ÍNDICE	VII
LISTA DE FIGURAS	IX
LISTA DE TABELAS	X
1. Introdução	1
1.1 Enquadramento	1
1.2 Motivação	1
1.3 Objectivos	2
1.4 Estrutura da Tese	2
2. Descrição do Problema	4
2.1 Introdução	4
2.2 Tipologias de imersão	6
2.3 A imersão como processo	8
2.4 Ambientes sonoros imersivos versus realidade virtual	10
3. Espaços sonoros	12
3.1 Soundscape	12
3.2 Organização dos espaços sonoros	15
3.2.1 Localização das fonte sonoras e reverberação	16
3.2.2 Espaço tangível	19
3.2.3 Espaço simbólico	21
3.2.4 Factor temporal	22
3.3 O ouvinte	23
3.3.1 Cognição	23
3.3.2 Realismo construído	24
3.4 Trabalho relacionado	27
3.4.1 Tapestry	27
3.4.2 Soundscape Modeling Technology	28
3.4.3 Interactive Soundscape designer	29

4. Projecto	31
4.1 Concepção	31
4.2 Estrutura do Algoritmo	32
4.2.1 Módulo de sampling	33
4.2.2 Módulo de decisão	33
4.2.3 Módulo de reprodução	34
4.2.4 Exemplo prático	35
4.3 Implementação	36
4.3.1 Ferramentas	37
4.3.1.1 Max MSP	37
4.3.1.2 KYMA	38
4.3.1.3 OpenSound Control	39
4.3.1.4 Processamento distribuído	40
4.3.2 Desenvolvimento da aplicação	41
4.3.2.1 Max/MSP - <i>Patch</i> principal	41
4.3.2.2 Max/MSP - Abstraction “Cenário”	43
4.3.2.3 Kyma - Som principal	45
4.3.2.4 Implementação do VCS (Virtual Control Surface)	47
5. Conclusões e Perspectivas de Desenvolvimento	48
6. Referências Bibliográficas	51
7. Anexo A	53

Lista de Figuras

FIGURA 1. PERCEÇÃO DA LOCALIZAÇÃO DAS FONTES SONORAS	16
FIGURA 2. RELAÇÃO ENTRE SOM DIRECTO E SOM REFLECTIDO.....	18
FIGURA 3. APLICAÇÃO TAPESTREA.....	28
FIGURA 4. SISTEMA GRÁFICO E CONCEITOS SONOROS.....	29
FIGURA 5. DIAGRAMA DA APLICAÇÃO SOUNDSCAPE MODELING.....	29
FIGURA 6. SOUNDSCAPE DESIGNER (APLICAÇÃO)	30
FIGURA 7. DIAGRAMA DO ALGORITMO.....	33
FIGURA 8. REPRESENTAÇÃO GRÁFICA DE UMA SUCESSÃO DE CENAS SONORAS	36
FIGURA 10. EXEMPLO DE UM PROTÓTIPO DO KYMA.....	39
FIGURA 11. PACARANA: O SISTEMA DE DSP MAIS RECENTE DA <i>SYMBOLIC SOUND</i>	39
FIGURA 12. <i>PATCH</i> PRINCIPAL DA APLICAÇÃO	42
FIGURA 13. ESTRUTURA DO FICHEIRO DE TEXTO.....	43
FIGURA 14. <i>MAX/MSP ABSTRACTION</i> “CENÁRIO”	43
FIGURA 15. ALGORITMO DE SELEÇÃO DO ID DO <i>SAMPLE</i> A SER TOCADO E ALGORITMO DE <i>DELAY</i>	44
FIGURA 16. TIPOLOGIA DE UMA MENSAGEM <i>OSC</i> : PREFIXO, PARÂMETRO, ID E VALOR.	45
FIGURA 17. FLUXO DE ÁUDIO E LIGAÇÃO DE PROTÓTIPOS NO SISTEMA KYMA	46
FIGURA 18. PARÂMETRO DE CONTROLO DE UM PROTÓTIPO DO KYMA ATRAVÉS DE <i>OSC</i>	46
FIGURA 19. <i>VCS (VIRTUAL CONTROL SURFACE)</i>	47

Lista de Tabelas

TABELA 1. RESUMO DA CLASSIFICAÇÃO DE SONS DE ROBERT MURRAY SCHAFER	14
TABELA 2. MANIPULAÇÕES SONORAS SEGUNDO UM MODELO DE OUVINTE (INGER EKMAN 2010)	26

Capítulo 1

1. Introdução

1.1 Enquadramento

A componente sonora, no âmbito das novas tecnologias, apesar de ainda ser, de certa forma, um campo de investigação e investimento menosprezado, tem conhecido recentemente novas oportunidades tecnológicas como o áudio processual, áudio interativo, etc.. que implicam novas reflexões, paradigmas e novos métodos exploratórios.

O trabalho descrito nesta tese tem como objecto de análise o conceito de imersão sonora no âmbito das novas tecnologias multimédia, fazendo um ponto da situação relativamente à pesquisa e conhecimento que existe, que tem vindo a ser produzido e que se pretende para o futuro no campo da criação e investigação sonora.

1.2 Motivação

A motivação principal para a realização deste trabalho, resulta do trabalho prático e profissional do autor e na necessidade sentida em sistematizar e refletir acerca da vertente sonora como elemento cada vez mais fundamental na elaboração de produtos multimédia. Neste sentido, foi notado que a imersão sonora, apesar de ser um tema susceptível de análise académica, é objecto de estudo sobretudo ao nível da reprodução dos seus conteúdos, nomeadamente a espacialização sonora. Ora, não tendo como objectivo fazer prova de um conceito que invalide o outro, pensámos que, a vertente ligada aos processos físicos de captação e reprodução sonora: Estéreo, binaural, surround, ambisónico, não explica totalmente a apreensão por parte do sujeito do ambiente sonoro, sendo que estes sistemas estão mais voltados para os aspectos técnicos, deixando por revelar um realismo muito particular, por vezes até construído, e que é âmbito de um imaginário no qual o *sound designer*¹, como criador, pelo processo de seleção, manipulação e mistura de sons é parte fundamental

¹ Desenho de som, design de som ou *sound design* é o processo técnico e criativo de manipulação,

no processo de decisão. Esta abordagem visa responder a questões mais do domínio psicológico e artístico que modulam a experiência imersiva inerentes à questão de como ouvimos, somos afectados pela manipulação de parâmetros do som ou como fazemos seleções e filtrações arbitrárias do conteúdo sonoro e que contribuem para níveis de imersão mais efetivos e relacionados com o ambiente visual e narrativo.

1.3 Objectivos

O objectivo deste trabalho consiste em estudar a imersão sonora mediante uma perspectiva teórica/prática. Esta pesquisa é centrada em contribuições de campos técnicos, científicos e artísticos diversos e analisa até que ponto estas contribuições poderão resultar numa abordagem híbrida que suporte e motive a construção de uma aplicação de *sound design* que agregue alguns destes conceitos.

Os objectivos específicos deste trabalho são os seguintes:

- Aplicação de teorias gerais sobre imersão à temática dos ambientes sonoros em conteúdos multimédia;
- Propor e desenvolver uma aplicação para *sound design* que opere sobre o conceito de *soundscape*², permitindo a geração e a recombinação da mesma em tempo real utilizando as plataformas de programação Max/MSP³ e KYMA⁴.

1.4 Estrutura da Tese

Este trabalho encontra-se estruturado em 4 capítulos dos quais, o primeiro é composto por esta introdução ao trabalho.

No segundo capítulo é apresentado e classificado o conceito e a questão central de partida da tese: o processo de imersão sonora e são discutidos paradigmas paralelos como o fenómeno perceptivo e cognitivo.

O terceiro capítulo é dedicado à discussão de algumas propostas existentes relativas ao processo imersivo sendo discutidos conceitos como o espaço acústico, a conceptualização de *soundscapes* e o papel do ouvinte em termos de *sound design*

² Conceito/ termo cunhado pelo compositor canadiano Raymond Murray Schafer (n. 1933), com base no termo *landscape* (paisagem) aplicado ao domínio sonoro no âmbito da sua investigação na área da ecologia acústica [*The New Soundscape* (1969), *The Turning of the World* (1977)]. [28]

³ www.cycling74.com

⁴ <http://www.symbolicsound.com/>

sugerindo que a realidade sonora dos espaços acústico é uma realidade manipulável e perceptiva.

Paralelamente a estes conceitos são discutidos modelos psicofísicos de espacialização e reverberação sonora que operam em simultâneo. Este capítulo pretende responder a sobre às questões: como e porquê a experiência imersiva não derivar somente do rigor e paralelo realístico, obtido através reprodução sonora mas também na sua conceptualização, produção, escolha e relação com pistas perceptivas e factores cognitivos que o conteúdo sonoro proporciona ao utilizador.

O quarto capítulo dedica-se à apresentação do trabalho prático proposto pelo autor que consiste na operacionalização de alguns conceitos inseridos na moldura teórica discutida nos capítulos anteriores. O projeto pretende descrever a concepção subjacente a uma ferramenta de *sound design* que recombina e sequencia, através de tabelas de transições, eventos sonoros usando métodos probabilísticos. Este método, permite a construção de cenários sonoros que podem ser totalmente aleatórios, ou obedecer a uma ordem sequencial decidida pelo autor.

Capítulo 2

2. Descrição do Problema

2.1 Introdução

Os *sound designers* confrontam-se com a necessidade de tornar a componente do som paulatinamente mais imersiva. No entanto, esta relação tem sido estudada de uma forma pouco aprofundada e pouca teoria ou prática tem sido desenvolvida e que estabelece uma relação explícita entre estas duas variáveis, apesar de ser amplamente reconhecida a importância do som para a imersão do sujeito. Esta situação, terá a ver sobretudo, com a evolução relativamente recente que a componente áudio tem verificado nos media digitais, derivada da optimização e incremento tecnológico ao nível do processamento e armazenamento e por alguma evidência científica e teórica de que o processo de imersão ocorre sobretudo a um nível inconsciente, tornando difícil definir um sistema que correlacione a imersão e o som [1].

Alguns da investigação em curso relativa ao tema examina o fenómeno através de uma perspectiva holística o que não fornece ao *sound designer* pistas suficientes para a compreensão do mesmo na sua relação com o áudio, do mesmo modo que outros paradigmas se centram somente em questões psicofísicas ou na temática da realidade virtual, não relacionando estes campos do saber com um modelo conceptual suficientemente abrangente e que inclua outras perspectivas como o envolvimento emocional, simbólico, cognitivo, etc., áreas igualmente relevantes e que, através do seu estudo, poderão ser alcançadas experiências imersivas mais eficazes.

Neste sentido, julgámos relevante a clarificação do conceito de imersão através da apresentação de abordagens diversas e que poderão servir como suporte teórico para a pergunta de base deste projeto, que consiste na análise do processo pelo qual a componente sonora contribui para a imersão do ouvinte e como esta componente é conceptualizada. A resposta a algumas das questões levantadas poderá servir de apoio para heurísticas futuras ou incorporação de novos modelos ou funções nas ferramentas de *sound design* existentes.

A definição de imersão levanta desde logo questões para qual as respostas são pouco consensuais na medida em que é um termo bastante abrangente, que se estende do domínio mais literal, como por exemplo um bebé imergido em líquido no útero materno, até ao efeito que decorre ao estar envolvido de uma forma muito profunda numa ação. A este propósito gostaríamos de citar as palavras de Joseph Nechvatal [2] quando diz se refere ao estado imersivo como:

“Immersion is the state of consciousness where an immersant's awareness of physical self is diminished or lost by being surrounded in an engrossing total environment; often artificial. This mental state is frequently accompanied with spatial excess, intense focus, a distorted sense of time, and effortless action”

Apesar de ser um termo amplamente utilizado, as definições são bastante escassas e há pouco consenso dentro da literatura sobre a definição de imersão [3]. Uma razão para isto, reside no facto de a imersão ser um termo que define um estado que é também relevante fora do contexto de meios digitais. Abrange o estado de estar submerso num líquido, bem como o estado de estar profundamente envolvido numa atividade, para dar os dois significados mais comuns, como já foi mencionado.

Muitas definições incluem propriedades pertencentes à experiência de imersão. Nachvidal [2] descrevem uma experiência imersiva como “aquela em que uma pessoa é envolvida por um sentimento de isolamento do mundo real”, que pode ocorrer tanto em filmes como em jogos. Um significado semelhante é encontrado em Dovey e Kennedy [4] Estes autores definem a imersão como “a experiência de perder um sentido de realização no presente, enquanto concentrado num ambiente mediado.”. No jogos, onde o jogador está a interagir, em vez de experienciar passivamente, o jogador diz-se perder de imediato o contacto com o ambiente físico. Estes autores descrevem a relação entre o jogador e o jogo para definir a imersão como “sensação do jogador de estar realmente no mundo do jogo.”

No contexto do cinema, o conceito de imersão também é usado. Carr [5] define imersão em relação à espacialização sonora: Esta acontece quando o som é percebido como vindo de todas as direções. Posteriormente, os ouvintes são imersos e “deixam de ser capazes de assumir a posição de um observador mais ou menos distante em relação ao objecto”. Por oposição, nos jogos, o utilizador não é um observador passivo, mas um participante ativo. À medida que o jogador está ativamente a interagir com o jogo, é mais fácil para este tornar-se absorvido pela atividade. Esta, é também por si só um aspecto da imersão.

Dovey e Kennedy [4] também afirmam que a imersão difere entre os media passivos (por exemplo filmes) e media interativos (por exemplo, jogos). Como o jogador é um participante ativo, diferentes efeitos emocionais relacionadas a esta participação são

implicados. Enquanto a "perda de sentido do eu", é provável que ocorra em muitas formas de consumo multimídia, os sentimentos de estar imerso no mundo do jogo em combinação com a intensa concentração é uma propriedade distinta da experiência de jogo.

2.2 Tipologias de imersão

No sentido de abordar o tema o problema, tem sido notada uma evolução na pesquisa acerca do conceito de imersão, começando pela tentativa de o classificar à medida em que vão sendo identificadas e discretizados diferentes tipos de imersão. Um denominador comum ao conceito parece ser a sua característica híbrida no sentido em que, agrega diferentes vectores, os quais podem operar, ou não, em simultâneo.

Algumas destas classificações irão ser abordadas e discutidas em seguida.

Segundo Taylor [6], convém distinguir dois tipos de imersão, apesar de ambos serem complementares. Por um lado, existe a imersão diegética, gerada pelo simples contacto com o conteúdo sonoro e com a consciência deste. Um exemplo disto são as primeiras reações aos estímulos visuais e sonoros provenientes de um conteúdo multimídia. Por outro lado, existe a imersão intra-diegética que implica a imersão do sujeito num espaço recriado através da perspectiva da personagem, incorporando o ponto de vista desta.

A perspectiva de este autor é bastante importante uma vez que, introduz o conceito de processo imersivo. Segundo o autor, o utilizador necessita de estar diegeticamente⁵ imerso para conseguir estar intra-diegeticamente imerso. Ou seja, introduz níveis de profundidade imersiva, ao mesmo tempo que divide o conceito em três aspectos básicos e de extrema importância para a análise e criação de ambientes imersivos: relação com a atividade do utilizador, o elemento espacial (presente no cenário da personagem) e o imaginativo (a perspectiva da personagem).

Ermi e Mayra [1], após investigarem e analisarem a experiência dos videojogos, sintetizam a sua investigação em três elementos que, segundo os autores são fundamentais no desenho e concepção dos videojogos: a) qualidade e estilo audiovisual b) nível de desafio c) mundo imaginário e fantasia. Estes elementos por sua vez encontram uma correspondência em dimensões imersivas: a) imersão sensorial b) imersão ligada ao desafio e superação c) imersão imaginativa.

⁵ Diegético: Conteúdo objectivo; todo o universo que é perceptível pelos personagens em cena na linguagem cinematográfica. Visivelmente identificável. [29]

Este modelo ficou conhecido através do acrónimo SCI (*Sensory, Challenge based, Imaginative*).

A primeira dimensão da imersão, a dimensão sensorial, concede uma relação com um mundo recriado através do qual o mundo real passa a ser o *background* deste, esta dimensão introduz o conceito de presença que irá ser abordado posteriormente e assume o mundo recriado como a nova realidade do sujeito, a sensação de “estar lá”.

O ênfase dado à qualidade sensorial prende-se com o facto de este se assumir como uma prova da experiência que vai potenciar os níveis de imersão seguintes, sobretudo o nível imaginativo.

Björk & Holopainen [7] definem os seguintes tipos de imersão acrescentando novas categorias que se podem revelar relevantes na construção de ambientes imersivos:

- Imersão espacial: ocorre quando o sujeito sente que o mundo construído é convincente do ponto de vista perceptivo na linha da proposta teórica de Ermi e Mayra [1]. O sujeito sente que está nesse mundo independentemente deste ser, ou não, uma emulação de uma determinada realidade.
- imersão emocional narrativa: o sujeito sente empatia com determinada personagem ou com estados emocionais desta. No âmbito sonoro, e a título de exemplo, esta pode ser obtida através de manipulação do áudio no sentido de transmitir determinados estados emocionais associados a determinada percepção sonora. Este tema irá ser desenvolvido no capítulo 3 e serão dados exemplos.
- imersão cognitiva: raciocínio abstracto, resolução de problemas. Implica níveis intelectuais mais profundos como o conhecimento anterior, a memória, etc..
- imersão sensório-motora: deriva do resultado de diversas ações repetitivas que influenciam as ações futuras do sujeito.
- Imersão psicológica: Segundo o autor, esta classificação deriva do facto de o sujeito confundir, a dada altura, o mundo real com um imaginário. A nosso ver, e no âmbito deste estudo, esta dimensão assume bastante relevo pela sua dimensão afectiva e através da qual são criados vínculos com os ambientes sonoros.

Calleja[8] desenvolve uma classificação bastante detalhada do conceito de imersão que introduz conceitos novos com alguma importância para a relação entre o áudio e estados imersivos. O autor centrou-se sobretudo na área dos jogos e denominou-os por envolvimento, relacionados com a incorporação do jogador no ambiente do jogo. Optámos por incluir o ponto de vista deste autor na medida em que inclui factores sócio culturais que servirão como ponto de partida para o tema da ecologia acústica e como esta poderá funcionar como factor de imersão.

À luz disto, o autor dividiu a sua classificação em seis itens:

- Envolvimento táctico: tomadas de decisão e planeamento.
- Envolvimento performativo: relacionado com o controlo em termos cinéticos, desde a aprendizagem destes até à sua execução e resposta internalizada.
- Envolvimento afectivo: Relacionado com todas as respostas cognitivas, emocionais e cinéticas entre o jogador e o ambiente do jogo.
- Envolvimento partilhado: Relacionado com o controlo de um avatar num ambiente representativo. A importância deste ponto reside na fixação da personagem em termos da sua localização, tanto ao nível espacial como social. Implica a temática da presença e de relação comunicacional com outras personagens do mundo do jogo ou da narrativa.
- Envolvimento narrativo: Relacionado com os elementos narrativos da história que pode envolver factos históricos e, descrições de ambientes e locais ou épocas específicas, descrição do cenário que suporte uma missão (narrativa designada) ou interpretação da mesma pelo jogador (narrativa pessoal).
- Envolvimento espacial: Relacionado com a localização em um determinado cenário que pode ir para além das fronteiras da imagem e fornecem pistas auditivas para elaborar mapas mentais que convidem à exploração do espaço e que alertem para potenciais ações na narrativa.

2.3 A imersão como processo

Como foi visto até aqui, a imersão é um produto multidimensional na medida em que envolve inúmeras variáveis. No entanto, achamos muito pertinente e adequada ao âmbito do conceito sonoro (temporal por natureza), a posição de Brown[9] que reflete acerca da imersão como processo temporal.

A densidade sonora pode evoluir ou diminuir à medida que o tempo vai passando e fornece elementos que potenciam a imersão do sujeito de uma forma gradual à medida que o enquadram progressivamente numa moldura e ambiente sónico. No âmbito do projeto proposto para a tese, esta questão assume bastante relevância uma vez que, a criação de ambientes sonoros imersivos, se pode debater com a questão oposta; a repetição de sons na janela temporal e/ou a monotonia dos mesmos pode interromper o processo imersivo e de credibilidade do ambiente sonoro. Ainda segundo Brown [9], é importante reconhecer que o som existe e se desenvolve no tempo. A imersão torna-se um processo de construção e não um estado que se relaciona diretamente com a atenção e concentração.

Neste sentido, definiram três estágios que definem o processo imersivo:

- Contacto
- Envolvimento
- Imersão plena

Esta abordagem teórica relativamente à imersão é de certa forma rara e original e contrasta com as abordagens anteriores uma vez que introduz a variável do processo temporal. Os autores salientam que até atingir o último estágio, o indivíduo terá que experienciar os anteriores e pode se deparar com factores disruptivos da imersão que condicionem o seu percurso até estarem plenamente imersos. Exemplos de factores que perturbem este processo estão, para o estágio do contacto, a necessidade em se concentrar e estar atento, para o estágio do envolvimento, o tipo de cenário é muito importante (visualmente e esteticamente) e para o estágio de imersão total consideram a atmosfera e a empatia com o cenário proposto como elementos preponderantes.

Com base nas definições, são distinguidos 3 aspectos fundamentais: transporte para o mundo virtual, absorção pela atividade e identificação e empatia com a situação ou ambiente.

Estas 3 dimensões podem ser ligadas à conceptualização do *design* de som e podem servir de como modelo teórico que suporte a nossa premissa, uma vez que poderão integrar as diversas variáveis que contribuem para o estado imersivo.

2.4 Ambientes sonoros imersivos versus realidade virtual

Pensámos, neste ponto após a apresentação e discussão das propostas teóricas acima que convém fazer uma distinção importante e que, de certa forma está na base do nosso projeto. A distinção entre realidade virtual e ambientes digitais imersivos:

Ambientes digitais imersivos podem ser entendidos como sinónimo de realidade virtual, não implicando, no entanto, que a “realidade” seja real mas, sobretudo, perceptiva. Um ambiente envolvente digital é sobretudo uma ponte entre o ouvinte e um universo de fantasia ou abstração, desde que o utilizador se sinta uma parte deste universo podendo ser ou não uma emulação da realidade [10].

Nas palavras de Durand R. Begault [11] *“This contrast underlies an important difference between the phrases **virtual environment** and **virtual reality** meaning that no correlation to reality is necessarily assumed grammatically by the first term, and we may certainly want to create virtual experiences that have no counterpart in reality. In fact, the phrase “naive realism” has been used to critique virtual environment designers that purport to effect perceptual experiences that are as real as those experienced under natural conditions.”*

Ou seja, se a qualidade da imersão poderá parecer uma variável difusa, não deixará de ser um critério avaliativo em termos de eficácia perceptiva e simulação. Neste sentido, a eficácia de um sistema de realidade virtual ou multimédia resultará na ligação deste com a qualidade de imersão. Segundo o mesmo autor, A simulação perceptual terá que ser vista à luz de factores humanos e baseados em investigação na área da psicofísica e cognição a fim de serem estabelecidas correlações entre a performance dos sistemas sonoros e os seus resultados na percepção do utilizador.

Isto é, supondo que a luta pelo realismo não é obrigatória, propõe-se um modelo que traduza a percepção do utilizador baseado na seleção de eventos sonoros, ou seu processamento e a sua ordem de reprodução sonora.

Após análise da problemática e do conceito teórico subjacente à imersão e a abordagens que a expliquem em termos sonoros e perceptivos, pretende-se uma aplicar, de uma forma prática, algumas conclusões destas abordagens através da concepção de um protótipo que permita a construção de ambientes sonoros.

Ainda à luz do que foi exposto acima e atendendo ao carácter híbrido do processo imersivo que parece integrar factores emotivos, cognitivos, psicofísicos e acústicos, a aposta numa aproximação ao processo através do recurso à manipulação sonora, resíntese e reprodução eventos baseada em processos semi-aleatórios e tendenciais utilizando o acesso a bases de dados e a relações probabilísticas entre eventos, pa-

rece-nos ser uma possibilidade válida que visa, entre outros objectivos a simplificação do processo de construção de *soundscapes* e o controlo sobre as mesmas.

Capítulo 3

3. Espaços sonoros

Após a introdução no capítulo anterior da problemática da imersão sonora, iremos centrar a investigação teórica na temática do espaço sonoro e a conceptualização do mesmo em termos de imersão do ouvinte. Serão discutidos alguns elementos e processos para a criação de ambientes sonoros imersivos e sugeridas hipóteses que sustentarão o projeto prático.

3.1 Soundscape

- O que é uma soundscape?
- *“An acoustic environment or an environment created by sound. The sonic environment. Technically, any portion of the sonic environment regarded as a field for study. The term may refer to actual environments, or to abstract constructions such as musical compositions and tape montages, particularly when considered as an environment.”* (R.M. Schafer, 1977:275) [12]

O termo *soundscape*, ou paisagem sonora, equivalente aural de uma paisagem natural, quase que automaticamente nos remete para o trabalho teórico de R. Murray Schafer, nomeadamente para o seu livro *The Tuning of the World (The Soundscape, 1977)* [12]. O autor conjuntamente com outros autores do *World Forum For Acoustic Ecology*, tais como Barry Trouax, entre outros, estudaram, de uma forma relacional sons, ambientes e culturas.

A terminologia relativamente ao conceito de soundscape tem vindo a crescer, fruto do interesse cada vez maior que vários campos de investigação têm demonstrado no tema, e o estudo do mesmo continua a desempenhar um função agregadora e trans-

versal, permitindo análises e taxonomias de paisagens já existentes ou recriação de novas paisagens, quer no campo científico, quer no campo criativo. [13]

Contudo, e apesar da proliferação de terminologia, de uma forma tradicional, a classificação de soundscapes inclui: *keynote sounds*, *signal sounds* e *soundmarks* de acordo com as suas funções socioculturais [12].

O método de análise de Schafer supõe como objecto de estudo paisagens sonoras do mundo real e, apesar de não mencionar a criação artificial de paisagens sonoras ou mundos sonoros imaginários ou de sequências temporais de eventos, o seu modelo de organização e classificação fornece uma moldura teórica que pode ser explorada em termos de *sound design* de uma forma ativa, uma vez que permite ao *sound designer* entender a relação de sons e a anatomia de uma paisagem sonora e construir, por exemplo, bibliotecas de sons e correlacionar a reprodução de eventos sonoros.

Neste sentido, o autor define os *keynote sounds* como sons ambiente constantes, mais subliminares, e que exigem esforço e concentração para serem identificados e entendidos (sons de tráfego sonoro ao longe); *signal sounds* que se caracterizam como sendo mais proeminentes e próximos de ouvinte, que os identifica de uma forma mais consciente (sinos de uma igreja, por exemplo) e os *soundmarks* que caracterizam um lugar [12].

Sons Naturais	Sons Humanos	Sons Sociais	Sons Mecânicos	Silêncio e Sossego	Sons como indicadores
Sons da criação	Sons da voz	Descrição geral de paisagens rurais	Máquinas		Sons de pratos e sinos
Sons do apocalipse	Sons do corpo	Paisagens sonoras urbanas	Equipamento industrial		Buzinas e assobios
Sons da água	Sons de roupas	Paisagens sonoras de cidades	Máquinas de transporte		Sons temporais
Sons do ar		Paisagens sonoras marítimas	Máquinas de guerra		Telefones
Sons da terra		Paisagens sonoras domésticas	Comboios e carruagens		Sistemas de alarme

Sons do fogo		Sons de comércio, profissões e estilos de vida	Motores de combustão		Sons de prazer
Sons dos pássaros		Sons industriais	Aviação, construção e equipamento de demolição		Indicadores de ocorrências futuras.
Sons dos animais		Sons de entretenimento	Ferramentas mecânicas		
Sons dos insectos		Música	Ventilação		
Sons de peixes e do mar		Cerimónias e festivais	Instrumentos de guerra e destruição		
Sons das estações		Parques e jardins	Maquinaria de trabalho de quinta		
		Cerimónias religiosas			

Tabela 1. Resumo da classificação de sons de Robert Murray Schaffer

Apesar de Schafer não se ter referido diretamente ao processo imersivo que ocorre para que um ouvinte sinta inserido dentro de um espaço acústico, sendo que este processo é objecto de análise nos pontos seguintes da dissertação, o seu modelo relacional menciona um processo dinâmico baseado em pistas perceptivas que recontextualizam permanentemente os locais em termos sonoros. O significado visual e acusmático⁶ é mediado pela ação física do ouvinte à medida em que se move no espaço e recategoriza perceptualmente os locais [14]. Esta característica ativa, traduz-se, recorrendo ao exemplo sonoro anterior, quando o ouvinte, ao continuar o seu trajeto, passa a igreja e se dirige ao local onde anteriormente ouvia o som ambiente difuso e constata que esse local é a praça central da cidade; esses sons de carros e buzinas deixam de ser *keynote sounds* e passam a ser *signal sounds*.

⁶ O som acusmático é o som que ouvimos sem que percebamos a fonte de origem, ou seja, algo que ouvimos sem saber de onde provem. [14]

3.2 Organização dos espaços sonoros

A percepção de espaços contempla a vertente sonora como indicador. As referências sonoras que um espaço transmite, o que comunica ao ouvinte, como se organiza auditivamente, ajudam à sua imersão na medida em que criam ambientes sonoros particulares.

Um ambiente sonoro compreende um cenário que pode ser visto à luz da experiência teatral como um repositório de componentes que se articulam e lhe dão vida através de um processo ativo que ocorre dentro deste mesmo espaço. Isto implica a compreensão do conceito de espaço acústico, decompondo alguns dos seus elementos e tentar explicar a sua importância no processo imersivo.

À luz da ideia de cenário sonoro, gostaríamos de introduzir os conceitos de espaço tangível e espaço simbólico explicando a sua relevância para a concepção de ambientes sonoros e articulando-os com a abordagem que pretendemos para o projeto prático proposto durante a tese. O conceito de espaço simbólico no âmbito da ecologia acústica integra elementos que fornecem características imersivas relacionadas com locais, períodos e factores de índole sócio cultural que ajudam o ouvinte a ser enquadrado sonoramente através de pistas perceptivas que o levam a assimilar e a recriar ambientes sonoros. Este conceito torna-se bastante importante à medida que irão ser discutidas as questões como o realismo construído ou emulação de realidades físicas.

O espaço tangível resulta, por outro lado, de propriedades sonoras que fornecem características espaciais e físicas como o volume, propagação sonora e reverberação.

A exploração do fenómeno sonoro para a percepção de espaços é bastante relevante no sentido em que acrescenta dimensões de análise para a conceptualização de ambientes acústicos. Esta análise ajuda a compreender como os utilizadores se sentem imersos dentro de um espaço. Factores como a localização e reverberação contribuem como ferramentas através das quais o *sound designer* conceptualiza e desenha os espaços em termos de propriedades tangíveis, integra os elementos simbólicos e a sequência de eventos sonoros. Pensamos que após estes temas terem sido discutidos, fará sentido tentar perceber a forma como estes elementos se interligam na conceptualização de ambientes sonoros com relevância cognitiva para o sujeito e como contribuem para a imersão deste no espaço acústico. [15]

3.2.1 Localização das fonte sonoras e reverberação

Difícilmente poderíamos descrever um espaço sem nos referirmos aos conceitos de localização e de reverberação. A articulação destas duas variáveis é fundamental para a percepção e caracterização acústica de espaços uma vez que fornecem pistas tangíveis relativas às dimensões do mesmo, por exemplo, e à distribuição dos eventos sonoros neste mesmo espaço.

A localização das fontes sonoras contribui de forma significativa para o enquadramento do ouvinte relativamente a um determinado espaço acústico através de indicadores de amplitude, tímbricos e de tempo que conferem ao espaço características tridimensionais derivadas de processos psicoacústicos e à forma como o cérebro soma e organiza a informação que chega aos dois ouvidos (fig.1). A localização sonora não é uma característica do som per se mas contribui para a percepção do mesmo, conferindo-lhe, pelas alterações tímbricas mencionadas aspectos qualitativos.[15]

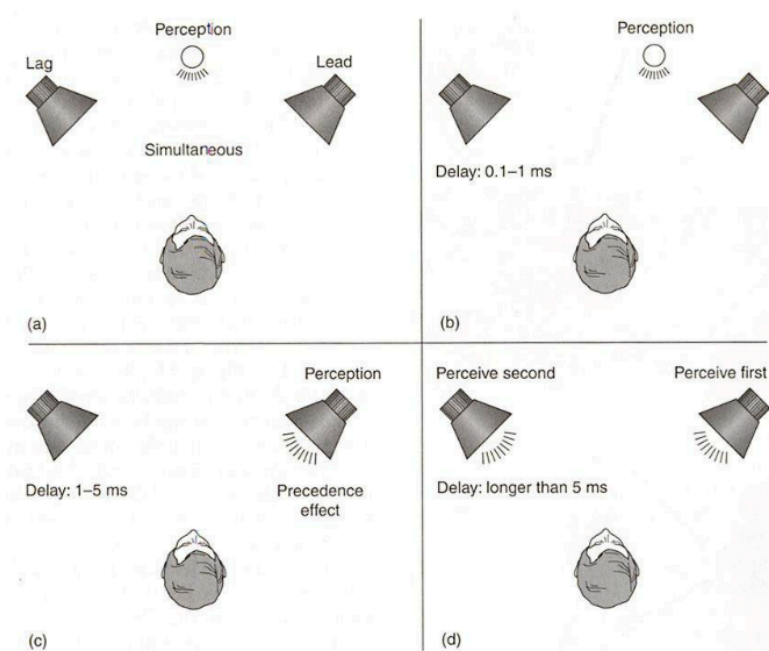


Figura 1. Percepção da localização das fontes sonoras

Ao defendermos que a imersão do ouvinte dentro de um espaço acústico está dependente da percepção de si próprio relativamente ao que o envolve, necessitamos de garantir que o ouvinte é capaz de discriminar a origem das fontes sonoras dentro do espaço acústico. Um ponto interessante salientado por LaBelle [16], vai ainda mais longe ao mencionar que, a importância em localizar as fontes sonoras, mesmo que

estas não sejam visíveis, reside nos indicadores que dão ao ouvinte acerca da existência de mais espaços envolventes paralelos, ao mesmo tempo que estimulam a sua imaginação e consequentemente os níveis de imersão narrativos e imaginários contextualizados no capítulo 2.

A ambiguidade relativamente à localização de determinados sons também poderá ser um instrumento que reforça a imersão sonora. É interessante este paralelo à luz das propostas de Schafer [12] referentes ao conteúdo espectral: segundo o autor, a ambiguidade na localização das fontes é frequência-dependente e sugere como exemplo as frequências mais graves. Estas frequências, são mais difíceis de localizar em termos de direccionalidade o que implica a percepção, por parte do ouvinte, de estar rodeado pelo som e, consequentemente, mais imerso neste espaço.

No âmbito da localização sonora, a variável frequência funciona, então, como um eixo, através do qual e sobre variações da mesma, a percepção da origem do som e da sua posição relativa é moldada e, consequentemente, a imersão imaginativa e perceptiva estimulada.

Como referimos anteriormente, nas linhas introdutórias a este ponto da tese, a localização das fontes sonoras per se, não sugere de uma forma direta, as dimensões de um espaço acústico, apesarem de sugerirem a forma como espaço soa em termos da sua “população” de eventos. A movimentação no espaço destes eventos e as variações de amplitude, densidade e intensidade dos mesmos relativamente a uma posição de referência, ajudam e sugerem estas hipóteses dimensionais, no entanto, para este efeito, a reverberação de um espaço proporciona os indicadores necessários de forma a aferir as suas dimensões, propriedades físicas dos materiais (madeira, metal, tecido, etc.), profundidade e distância à fonte sonora.

Um espaço reverberante fornece pistas psicoacústicas ao ouvinte relativas à sua distância a uma determinada fonte sonora através de uma taxa entre o som direto e o som reflectido, sendo o primeiro todo o som que chega ao ouvinte sem qualquer reflexão nas superfícies que delimitam o espaço e o segundo todo o som reflectido por estas mesmas superfícies (Fig. 2). Deste modo, quanto maior for a distância a uma dada fonte sonora, menor é a diferença entre estas duas variáveis e maior a sua equivalência em termos de intensidade. [17]

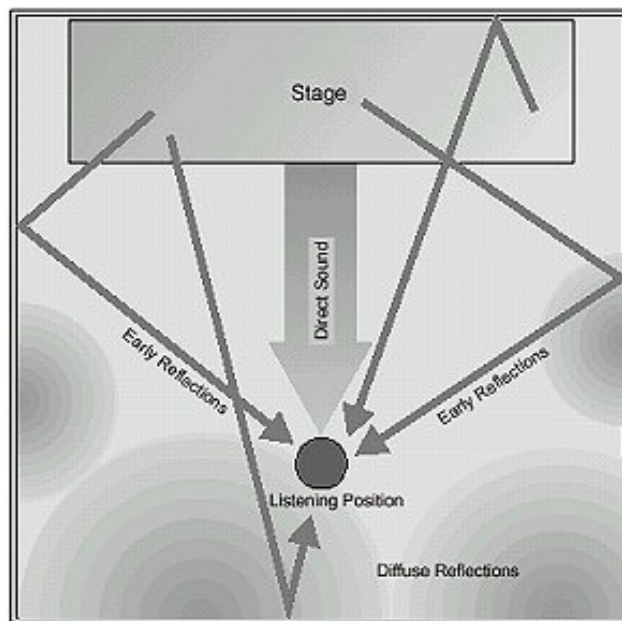


Figura 2. Relação entre som directo e som reflectido.

O factor reverberante contribui, então, para a vertente dimensional dos espaços e ajuda a contextualizar os mesmo integrando a localização das fontes sonoras ao mesmo tempo que acrescenta elementos relativos às propriedades dos materiais constantes nesse espaço. Outra das funções desta variável, refere-se às suas propriedades temporais o que implica que tempos de reverberação mais longos pressupõem espaços com dimensões maiores ou com características menos absorventes, do mesmo modo que, tempos de reverberação mais curtos, sugerem espaços de dimensões mais reduzidas, ou com propriedades materiais bastante absorventes. [17]

As propriedades materiais têm também como resultado audível, alterações ao nível tímbrico nas reflexões acústicas de um determinado som, pelo que caracterizam sonoramente um espaço, não só em termos quantitativos, isto é, tempos de reverberação maiores ou menores, mas sobretudo em termos qualitativos: reflexões timbricamente mais ou menos brilhantes (usámos o adjetivo brilhante no sentido sinestésico que se tornou comum como forma de descrição de conteúdo áudio onde predominam as frequências mais altas no espectro).

No sentido do que propomos para o trabalho da tese, os dois factores (localização e reverberação) surgem como elementos bastante importantes na medida em que contextualizam os elementos tangíveis e sócio culturais que compõem os cenários sonoros ao conferirem a estes espaços propriedades próprias que permitem ao ouvinte identificar o espaço e atribuir-lhe significados, estimulando desta forma a imersão perceptual e cognitiva.

Shinn-Cunningham e Schwartz [18] sublinham o papel da reverberação como ferramenta que potencia o realismo subjetivo de um cenário sonoro virtual. Este ponto de vista, apesar de conduzido por um estudo científico objectivo, levantou duas questões interessantes para a criação de ambientes imersivos: a) estabelece uma dicotomia entre o rigor da percepção da distância e a origem topográfica da fonte sonora. Se por um lado os dois factores não são independentes e o aumento de energia reverberante degrada os indicadores relativos à origem da fonte, por outro esta energia reverberante é necessária de forma a conferir ao espaço acústico um factor realista e b) sugere que o processo imersivo beneficia mais do compromisso entre estas duas variáveis, ainda que o rigor da localização das fontes seja, de certa forma, comprometido.

Este estudo parece indicar que a imersão sonora está correlacionada com algum índice de subjetividade. Se, à primeira vista, parece sugerir um paradoxo relativamente à importância atribuída à localização rigorosa de fontes sonoras, lembrámo-nos a conclusão dos estudos de Schafer quando menciona a ambiguidade introduzida, neste caso pelo conteúdo espectral mais grave e as evidências que sugerem que esta ambiguidade é muito importante para o processo imersivo.

No mesmo sentido, a dose de subjetividade sugerida por estas pesquisas, poderá revelar um mecanismo de desconstrução da percepção e cognição humana relativamente ao processo imersivo e constituir-se como uma ferramenta de *sound design* conduzida por realismos construídos e que dá ênfase ao papel ativo do ouvinte.

3.2.2 Espaço tangível

Um espaço tangível é um vector dimensional no sentido de, e para o propósito da tese, se constituir como uma classe que contém propriedades tangíveis como o volume e tempo sendo que estas propriedades são aferidas através de pistas perceptivas como a localização e a reverberação. Estes factores, sobretudo a interação entre eles proporcionam a recriação de ambientes virtuais e a imersão do ouvinte neles.

Criar estes ambientes implica, portanto, integrar elementos que derivam dos factores localização e reverberação, fazendo variar pistas perceptivas que indicam profundidade e tamanho do espaço, fazendo variar o conteúdo espectral dos elementos, a relação entre som direto e o som reflectido, a movimentação dos elementos no espaço, a sua velocidade de movimentação e a sua amplitude relativa.

Estes elementos, ao operarem em conjunto, sugerem ao ouvinte que se encontra dentro de um determinado ambiente acústico com limites identificáveis onde o cenário sonoro se desenvolve. Certos autores, como Chion [14], sugerem, a este respeito, que a percepção deste espaço ressonante é uma premissa para questões como o som

diegético⁷/não diegético⁸. Ao delimitarem um espaço, funcionam como uma âncora acústica que prende o ouvinte e contextualizam um determinado espaço acústico, ao mesmo tempo que permitem que novas pistas perceptivas sejam introduzidas, sugerindo a existência de mais espaços envolventes.

O espaço tangível é, então, um recipiente de pistas perceptivas que poderão complementar a classificação dos eventos sonoros que aí ocorrem. A título de exemplo poderemos imaginar um cenário físico como um café: neste cenário, podemos encontrar uma panóplia imensa de eventos sonoros: vozes, ruído de várias máquinas, ruído exterior, eventualmente louça a partir. O espaço tangível indica ao ouvinte um espaço físico tal como as dimensões do café, o tipo de superfícies, se, eventualmente, existem pisos adjacentes ou não, que as pessoas entram e saem, que se movem através de alterações tímbricas etc.. Mas, apesar de podermos reconhecer que se trata de um espaço social como um café, pelas diversas classes de som que o compõe, esta dimensão referente ao local e à sua natureza social é motivo de análise no item seguinte relativo aos espaços simbólicos. Esta dimensão física dos espaços, permite que os eventos sonoros “soem” de uma determinada maneira, agregando-os às características comuns do espaço e proporcionando-lhes um contexto e comportamento acústico específico através da manipulação das suas características tangíveis.

Nas palavras de Grimshaw [15], estas propriedades tangíveis funcionam como a porta de entrada para a criação de ambientes sonoros virtuais no sentido de sugerirem ao ouvinte a sua inclusão nesse espaço delimitando-o fisicamente. Após esta fase o ouvinte poderá ser estimulado a ser imerso imaginativamente através de técnicas de manipulação das pistas perceptivas que vão decorrendo no espaço.

A relação do ouvinte com estas transformações sónicas poderão ser casuais no sentido de não necessitarem, obrigatoriamente de seguir uma determinada ordem, pelo que as alterações que nele decorrem implicam mudanças perceptivas. A manipulação destas variáveis, ou pelo menos dos seus parâmetros, será fundamental no desenho de som destes espaços uma vez que será, através deste recurso, que poderemos argumentar se os níveis de imersão poderão ter uma relação direta com a realidade perceptiva do ouvinte. Uma relação inclusivamente mais efetiva do que através da emulação da realidade.

Estas alterações nas relações entre estes parâmetros do espaço poderão, então, implicar atribuições de significados diferenciados ao som, criando personagens ou ouvintes imaginários com estados perceptivos e cognitivos diferentes. Imaginemos duas

⁷ Som Diegético: sonoridades objectivas; todo o universo sonoro que é perceptível pelos personagens dentro de uma cena. [14]

⁸ Som Não-Diegético: sonoridades subjetivas; todo o universo sonoro que não é perceptível pelos personagens em cena. [14]

situações distintas: Uma personagem sentada na mesa de um café com vozes e diversos ruídos do ambiente. Ao variarmos a percentagem de som direto e de som refletido, ou ao variarmos a localização das fontes sonoras ou o seu timbre, teremos implicações diversas na representação cognitiva da personagem e do cenário acústico: Poderá a estar a prestar atenção a uma conversa em particular (o efeito de *cocktail party*⁹), poderá ser a representação de uma voz interior por contraste com o que se passa no espaço acústico, poderá tornar o som difuso e menos inteligível, o que poderá significar confusão mental ou pura e simplesmente que o café tem uma acústica péssima, o que cria um cenário tenso, denso e confuso.

Esta categorização reforça a ideia de um realismo perceptivo diverso e construído através da imersão imaginativa do ouvinte no interior de um espaço identificativo, que o fazem elaborar associações a determinados estados e mudanças neste espaço.

Em jeito de conclusão poderíamos resumir o que foi dito em relação ao conceito de espaços tangíveis, como sendo uma dimensão física, através da qual o ouvinte identifica um espaço em termos das suas propriedades quantificáveis e se contextualiza experienciando a imersão sensorial. Através da manipulação destas variáveis derivadas dos conceitos de localização e a reverberação, o ouvinte ao atribuir significados a estas repercussões psicofísicas poderá reforçar a imersão perceptiva e imaginativa. Estas questões serão debatidas no ponto 3.3 dedicado ao papel do ouvinte e na conceptualização deste por parte do *sound designer*.

Por sua vez, os espaços simbólicos, que irão ser discutidos em seguida, definem os espaços em termos de locais ou interações sociais específicas.

3.2.3 Espaço simbólico

Como foi referido nas notas introdutórias a este capítulo, o espaço sonoro, no âmbito da estrutura teórica que considerámos adequada para a tese, compreende espaços tangíveis e espaços simbólicos. Se os primeiros definem e caracterizam dimensões físicas, os segundos dizem respeito a categorizações locais específicas e compreendem elementos que reconstroem os espaços com base em conhecimento anterior e pistas provenientes da aprendizagem social. [12]

Esta caracterização dos espaços acústicos através de factores simbólicos, usa elementos identificativos dos locais, como por exemplo: sons de sinos de igreja conjun-

⁹ O efeito *cocktail party* (também conhecido como atenção seletiva) é um fenómeno psico-acústico que consiste em ser capaz de centrar a atenção num determinado estímulo sonoro filtrando-o de outros estímulos paralelos. [17]

tamente com passos em pátios largos e vozes de pessoas enquanto esporadicamente se ouve o som de carroças e de cavalos.

Na cultura ocidental a interpretação remeteria, talvez, para um tipo de cerimónia religiosa numa qualquer aldeia remota, talvez *Amish* pela ausência de automóveis, ou então para um passado remoto, ou, dependendo do volume das vozes ou tipo de toque do sino, para uma cerimónia mais específica como uma missa, baptismo ou funeral.

Poderemos sugerir que, se os espaços simbólicos classificam perceptualmente e cognitivamente o que se ouve, os espaços tangíveis definem a maneira como o espaço faz soar os elementos simbólicos do espaço físico.

Em termos de desenho de som, a importância deste elemento reside no potencial que tem para constituir cenários sonoros no qual interagem sons que caracterizam o ambiente não só pelas sua classificação (animais, máquinas, vozes, natureza, etc..), mas também pela sua frequência temporal (quando ocorrem, se são constantes ou esporádicos) e densidade (maior ou menor número de elementos que acontecem em simultâneo). Esta operação em conjunto, permite ao ouvinte atribuir conotações através de referências cognitivas identificando o espaço de uma forma mais complexa, atribuindo-lhe características próprias tais como o período histórico, por exemplo, ou adjectivá-lo como espaço movimentado, calmo, tenso, urbano, rural, etc..

3.2.4 Factor temporal

O fenómeno sonoro é intrinsecamente temporal, no entanto este termo pode ter significações perceptuais e cognitivas distintas.

Na concepção de paisagens sonoras e recorrendo a exemplos de cenários sonoros possíveis à semelhança do que foi feito no tema anterior relativo aos espaços simbólicos, o factor temporal pode significar: uma época e evento específico (guerra dos 100 anos através da predominância de sons de armas e utensílios de época), representação de uma janela temporal objectiva (o som de um relógio a indicar as horas), o volume de um espaço através do tempo de decaimento da sua reverberação (som de um tiro), um local específico a uma determinada altura do dia (som de rio e de insectos noturnos), etc.. ou a combinação de todos através de uma forma sequencial.

A reflexão sobre este elemento através do exercício imaginativo do parágrafo anterior pretende clarificar a ambiguidade do conceito através de alguns exemplos práticos e servir também como uma porta de entrada para a classificação das categorias temporais.

No âmbito da tese e do projeto proposto, este factor é um recurso conceptual que pode incluir uma sequência ao nível da linguagem cinematográfica, obedecendo a critérios probabilísticos determinados pelo *sound designer*. Neste sentido, um determinado tipo de evento aumenta a probabilidade de outro evento específico ocorrer. Um exemplo concreto será uma sequência imaginária durante a qual acontece um som de um tiro e a probabilidade de o evento seguinte serem sons de pássaros a voar é mais elevada.

Com recurso à manipulação e recombinação das representações temporais mencionadas, está no âmbito do controlo do *sound designer*, determinar o desenrolar do cenário sonoro através de uma perspectiva narrativa, estimulando, desta forma, a imersão perceptiva e cognitiva do ouvinte.

3.3 O ouvinte

O conceito de imersão em termos de processos espaciais e temporais proposto nos subcapítulos anteriores, diz respeito a características, cujas transformações estão na base dos julgamentos perceptivos e cognitivos do ouvinte. No entanto, o seu objectivo principal foi a constituição de uma moldura teórica que poderá permitir a conceptualização de espaços sonoros com ênfase na sua caracterização, sendo que o papel do ouvinte emanou de uma forma indireta.

Seguindo a mesma lógica subjacente aos espaços sonoros, subdividindo-o em elementos interligados mas independentes, parece-nos fundamental, neste tópico, expor o papel ativo de ouvinte no processo imersivo e de que forma este processo, derivado de julgamentos perceptivos, poderá implicar uma realidade construída.

3.3.1 Cognição

A cognição refere-se a um ramo derivado da psicologia perceptiva de alto nível, no sentido de implicar processos mentais mais complexos do que a simples sensação de estímulos tais como a memória, compreensão e tomada de decisão [19].

A complexidade e subjetividade implícita neste factor, revela que, questões como a familiaridade e a expectativa introduzem modificações no julgamento de determinadas pistas perceptivas, tornando a análise desta dimensão mais difícil.

Este subcapítulo, dedica-se então questionar um modelo de ouvinte fazendo a ressalva para o facto de, apesar do grau de subjetividade envolvido nestes paradigmas, o potencial criativo que possuem e, mais do que isso, a necessidade de implicar estas

questões ao investigar acerca da eficácia da imersão de determinados espaços sonoros.

3.3.2 Realismo construído

O conceito de cognição, como foi referido acima, introduz novas variáveis e um grau de subjetividade tão elevado que nos poderíamos questionar se o realismo sonoro, no contexto dos media, é um factor fundamental para a imersão do sujeito ou se, pura e simplesmente depende do conteúdo e do objectivo.

Segundo vários autores [14][20], para objectivos imersivos, um realismo reduzido poderá ser suficiente. Ainda nas palavras de Chion [14], “um realismo não inteiramente real”.

Estas perspectivas oferecem várias oportunidades não só de análise, mas também práticas na construção de novos métodos para *sound design* uma vez que poderemos sacrificar, até certo ponto, a vertente realística em prol de métodos que, através da imersão imaginativa e sensorial cumpram o objectivo de imergir o sujeito no interior de ambientes sonoros virtuais.

Esta abordagem não deixa, contudo, de apresentar à primeira vista uma questão paradoxal uma vez que, apesar de propor uma realidade que não pretende sempre emular uma realidade física, utiliza fenómenos psicoacústicos que simulam situações reais. Segundo Chion [14], o objectivo é transmitir uma impressão realista ainda que esta não tenha uma correspondência com nenhuma realidade em concreto.

Recentemente este paradigma tem produzido, ou pelo menos sugerido, uma motivação crescente para o estudo do papel ativo do ouvinte, não o considerando como um mero recipiente de estímulos sonoros. Estas abordagens centram a atenção em parâmetros das ciências cognitivas tais como as emoções, fenómenos como a atenção seletiva e as distorções perceptivas derivadas de situações de stress ou traumáticas, para produzirem conhecimento que eventualmente resultará num modelo de ouvinte.

Teóricos como Lastra [20] e Ekman [21] reforçam esta visão de realidade construída, talvez melhor denominadas como realidades, uma vez que propõe, para o mesmo ambiente sonoro apreensões perceptivas diversas derivadas da manipulação sonora. Os estudos destes autores sugerem que a relação com a realidade é até certo ponto simbólica, ligada a construções mentais que os sujeitos elaboram sobre os objetos sonoros mais do que os objetos sonoros em si e a sua fidelidade com os sons reais. Neste sentido, Van Leeuwen [22] exemplifica a utilização de recursos sonoros como o exagero da amplitude e frequências graves que tornam certos ambientes sonoros capazes de induzir reações viscerais, potenciando a imersão sensorial, ao mesmo tempo que sacrificam a vertente realista em prol do impacto emotivo.

Na mesma linha de pensamento, Ekman [21] vai mais longe ao propor um modelo de ouvinte subjetivo que implica uma conceptualização interpretativa de estados psicológicos que o autor dos cenários sonoros pretende transmitir. Segundo o mesmo autor, estados psicológicos como as intenções, emoções e pensamentos influenciam a audição e derivam de manipulações sonoras conforme a tabela 2.

Categoria do efeito	Manipulação	Motivação	Descrição	Exemplo
Seleção de sons	Controlo da reprodução, prioridades entre sons	Atenção, percepção multimodal	Escolha de sons de acordo com centros de atenção. Dependendo do campo de visão dar prioridade a certos sons, abnegando outros.	Dentro de um cenário visual denso, reproduzir somente sons relacionados com a atividade da personagem, sugerindo a sua concentração.
Controlo de volume	Mistura	Atenção, distorções perceptivas	Os sons relativos dos diversos sons serem alterados de forma a sugerir impactos emotivos ou processos de atenção. Sugere subjetividade pelo silêncio ou intensidade	Aumento do volume do som ambiente e baixando o som da personagem (passos, manuseamento de ferramentas sugerindo um ambiente opressivo face à personagem).
Transformação	Filtros	Distorções perceptivas	Pode significar estados alterados, emocionalidade extrema, stress, audição afectada.	Filtros passa baixos aplicados aos sons ambiente sugerindo que a personagem está afectada fisicamente ou mentalmente
Sons adicionais	Controlo da reprodução	Sons do corpo, alucinação	Introdução de sons do corpo como respiração ofegante	Selecionar sons de vozes e sussurros de uma forma caótica

			te, batimentos do coração, diálogo interno para reforçar o impacto emocional. Alucinações, memórias.	tica e espacialização aleatória sugerindo a perda de discernimento e estado de confusão.
Localização sonora	Mudanças na origem posicional da fonte sonora, concentrar todo o som em pontos únicos	Distorções perceptivas e de atenção	Agrupamento de objetos numa localização consoante a sua relevância em termos de atenção. Impacto emocional ao impedir a estabilidade da fonte sonora.	Mover abruptamente sons ambiente provocando um sentido de desorientação e surpresa.
Espacialização sonora	Acústica virtual	Atenção, alucinação. distorção perceptiva	Reverberação e eco para similar estados alterados	Contraste entre reverberação exagerada e sem reverberação para significar sons no interior da mente.

Tabela 2. Manipulações sonoras segundo um modelo de ouvinte (Inger Ekman 2010)

Este modelo pode ou não pretender uma representação natural dos ambientes sonoros mas cumpre, sobretudo, uma estratégia expressiva e narrativa.

Do nosso ponto de vista, este paradigma centrado na perspectiva do ouvinte, poderá constituir-se como uma importante moldura conceptual uma vez que transmite ao autor dos cenários sonoros um papel de decisão de quando e como deve escolher, manipular e reproduzir os sons de forma a criar o ambiente psicológico que pretende.

A manipulação de sons e o controlo independente dos elementos dos ambientes sonoros permite influenciar a imersão sonora através de diversos tipos de modificação nas relações entre os sons enfatizando alguns elementos e orientando a atenção do ouvinte intencionalmente para determinados focos sonoros.

Em termos práticos de *sound design* e na conceptualização de espaços acústicos através das perspectivas apresentadas, poderemos ter a opção adicional de optar por uma perspectiva que não privilegia, ou acha fundamental a emulação de realidades

concretas, mas que pretende criar cenários sonoros mais expressivos através da prioridade que dá a alguns eventos sonoros e à forma como os processa em termos de efeitos sonoros e os equilibra em termos de mistura.

3.4 Trabalho relacionado

Esta parte da dissertação pretende dar conta de alguns projetos práticos que abordam a temática da construção de *soundscales*. Não pretendemos referir exhaustivamente todos os projetos relacionados e por isso o critério subjacente à nossa escolha teve como base dois critérios:

- a) Dar a conhecer um pouco do tipo de trabalho que se faz nesta área mostrando perspectivas diferentes. Para isso selecionámos alguns das aplicações mais mencionadas e consideradas pela comunidade. Isto foi aferido através da pesquisa por diversos fóruns e *browsers online*.
- b) Mostrar, por contraste, que a nossa abordagem opta por um caminho diferente que poderá, no entanto, ser complementar.

3.4.1 TapeSTrea¹⁰

TAPESTREA : Techniques And Paradigms for Expressive Synthesis, Transformation and Rendering of Environmental Audio (also known as taps).

Este projeto foi desenvolvido por Ananya Misra, Perry R. Cook, Ge Wang no departamento de Ciências da computação e música da Universidade de Princeton, com o objectivo criar uma aplicação orientada para a transformação de cenários sonoros através de métodos de análise, síntese e manipulação de sons complexos (fig. 3).

Tomando como ponto de partida múltiplas gravações independentes, o utilizador pode extrair das mesmas elementos individuais que considere necessários para a composição de novas composições sonoras.

A aplicação utiliza diversos tipos de síntese tais como: síntese concatenativa¹¹ e síntese subtrativa¹² o que, segundo os autores, permite a manipulação de elementos

¹⁰ <http://taps.cs.princeton.edu/>

¹¹ Técnica de síntese também denominada por mosaico baseada na extração e reordenamento de pequenos *samples* de som [33].

¹² Técnica de síntese que consiste em filtrar elementos de espectros sonoros ricos harmonicamente [34]

individuais de uma gravação independentemente de sons de fundo e outros eventos sonoros que estejam a ocorrer em simultâneo obtendo-se resultados perceptivos muito convincentes.

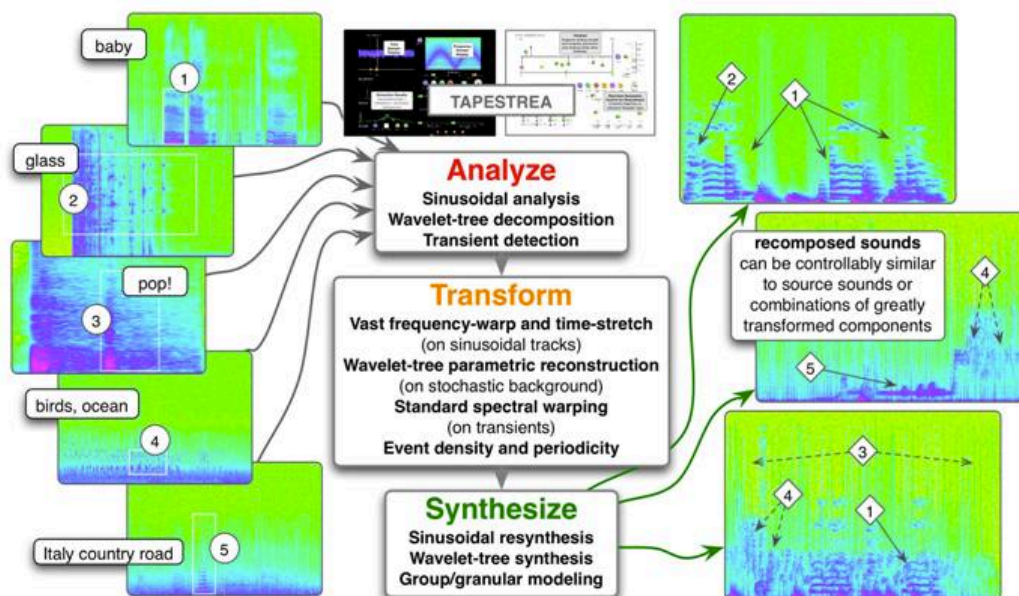


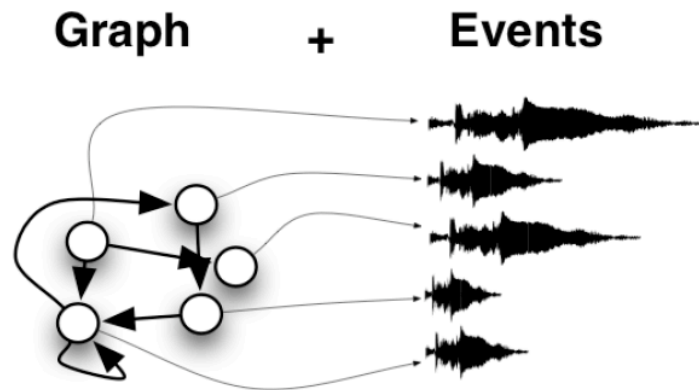
Figura 3. Aplicação TapeSTREA

Através de um ambiente gráfico ou através de *scripts* na linguagem de programação *Chuck*, o utilizador pode reorganizar os elementos que extraiu espacialmente e temporalmente e ir, continuamente, montando o cenário que pretende.

3.4.2 Soundscape Modeling Technology¹³

Este sistema foi desenvolvido pela Universidade Pompeu Fabra pelo Music Technology Group em Barcelona com o intuito de simplificar o processo de *authoring* e focalizado em realidade virtual e aumentada. É um sistema generativo que ambiciona a construção de paisagens sonoras realistas e interativas através de uma plataforma colaborativa (www.freesound.org) e de um motor de busca de elementos sonoros com recurso a *tags* e através de um algoritmo de síntese baseado em modelos gráficos (Fig. 4). Estes *samples* podem ser obtidos através do acesso a um repositório de áudio através do algoritmo de síntese que está alojado num servidor. O servidor gere os pedidos dos utilizadores e da aplicação de *authoring* desenvolvida em *SuperCollider* devolvendo a paisagem sonora aos utilizados construída através da rede (fig. 5).

¹³ <http://mtg.upf.edu/technologies/soundscapes>



The sound concept

Figura 4. Sistema gráfico e conceitos sonoros.

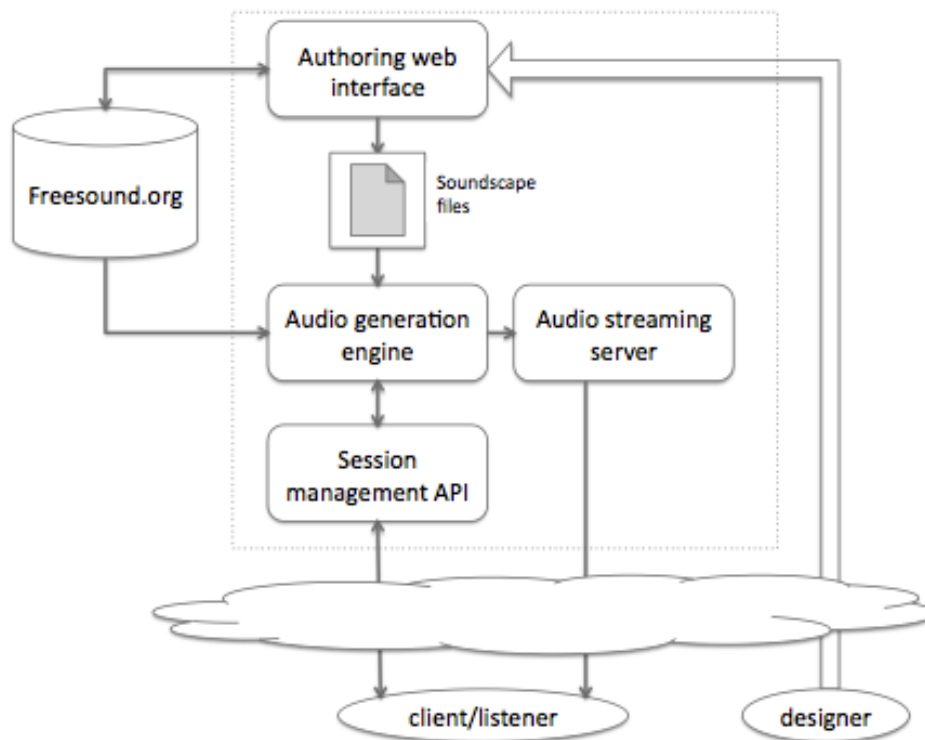


Figura 5. Diagrama da aplicação Soundscape Modeling

3.4.3 Interactive Soundscape designer¹⁴

Este software foi criado por Zach Poff e N.B. Aldrich em 2007 com o objectivo de proporcionar uma plataforma colaborativa inserida dentro de um ciclo de instalações

¹⁴ <http://www.interactivesoundscapes.org/>

sonoras em Portland. A aplicação é, nas palavras dos autores, uma oportunidade para as comunidades desenvolverem um perfil ou autorretrato utilizando recombinações de gravações áudio.

O utilizador pode povoar a superfície a negro (Fig. 6) com gravações de ambientes sonoros, decidindo quanto ao seu posicionamento no espaço e quanto às relações de amplitude entre os sons movendo os círculos coloridos nos planos x e y. Pode posteriormente gravar os novos cenários obtidos a fim de os visitar no futuro ou serem reutilizados por outros utilizadores da comunidade.

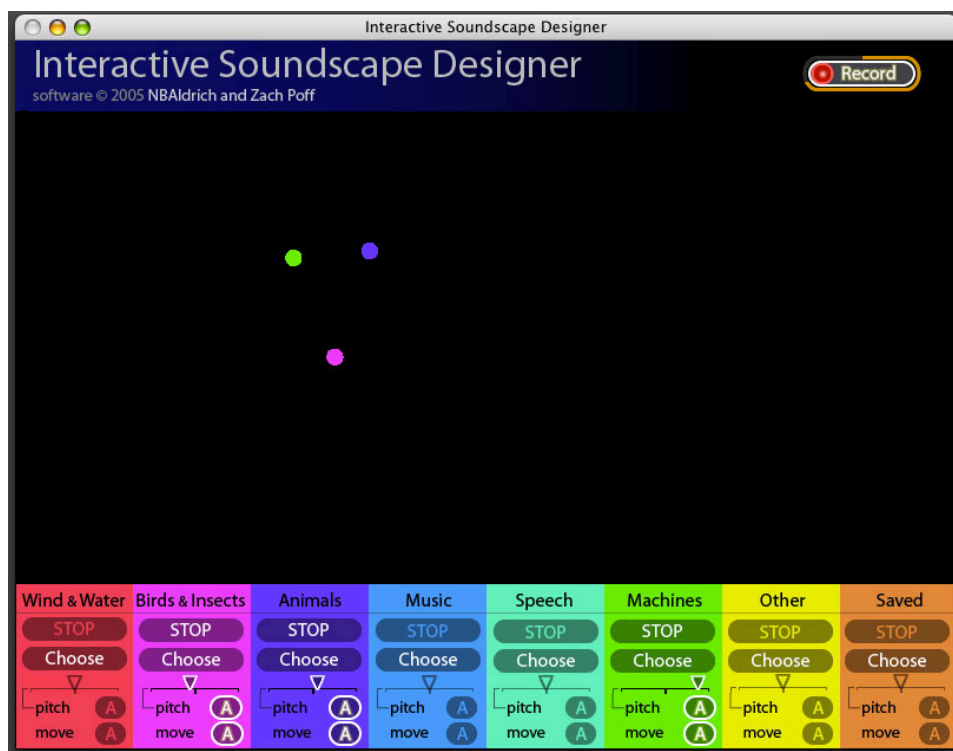


Figura 6. Soundscape designer (aplicação)

Capítulo 4

4. Projecto

4.1 Conceção

“The notion of soundscape is increasingly relevant not only in contemporary culture and acoustic ecology but also in many other fields, such as, e.g. audiovisual productions, in which soundscape is now a fundamental part of the whole soundtrack”.

Andrea Valle, Vincenzo Lombardo, and Mattia Schirosa

Nos capítulos anteriores foram discutidas algumas conjecturas teóricas que tentam explicar o processo imersivo introduzindo diversas variáveis e sugerindo um modelo de realidade perceptiva que, a nosso ver, poderá inspirar novos modelos de geração de ambientes sonoros.

A concepção do projeto teve como base a exploração de cenários acústicos através da construção de um algoritmo que permita a geração de eventos sonoros de uma forma sequencial e automática baseados na análise de contextos acústicos reais ou imaginários.

Esta metodologia organiza a sequência de eventos seguindo uma lógica de guião sonoro através do qual o autor determina probabilisticamente as relações sequenciais e temporais entre uma biblioteca de sons previamente construída e armazenada numa base de dados utilizando uma tabela de transições que define estados entre eventos. Isto permite um processo generativo capaz de criar um número infinito de cenários a partir de um número finito de sons (a lista de sons selecionados pelo autor).

Este processo cria e recombina as relações entre os sons gerando continuamente paisagens sonoras que não se repetem e sobre as quais o autor tem algum controlo, não só ao elaborar a potencial sequência aleatória, através de uma matriz, mas intervindo ao nível de parâmetros sonoros que agem sobre os eventos gerados tais como a sua amplitude, altura, velocidade de reprodução, panorâmica e relação entre o som direto e o som reflectido.

À luz de algum do trabalho relacionado referido no capítulo anterior, estamos em crer que esta abordagem apresenta alguma inovação no sentido de tentar intervir mais ao nível da imersão imaginativa e sensorial devido às possibilidades de manipulação sonora e de se enquadrar sobretudo dentro do domínio criativo. Neste sentido, poderemos dizer que a intenção de uma reprodução realista de cenários sonoros não é a principal premissa do projeto ainda que exista essa possibilidade dependendo dos objectivos do autor.

Os projetos relacionados que considerámos relevantes apresentar, têm diferenças quer ao nível do método de geração, quer de manipulação. Síntese “concatenativa” e análise espectral no caso do TapeSTREA e do MTG’s modelling Technology e baseado em *loops*, no caso do Soundscape designer. Em todos os projetos não existe uma orientação narrativa, daí pensarmos que a nossa contribuição assenta, sobretudo, neste factor.

4.2 Estrutura do Algoritmo

O algoritmo funciona de uma forma cíclica no sentido em que, dispara *samples* sonoros que por sua vez disparam outros *samples* sonoros, realimentando e recondicionando o processo conforme o diagrama que consta na figura 7.

O módulo de decisão, após o *input* do utilizador que inicia o processo manualmente, começa a rotina de leitura de uma tabela de transições que determina, para cada *sample* de uma dada *playlist*¹⁵ (compilada pelo autor do cenário), um valor mínimo e um valor máximo em unidades de tempo (t_{min} t_{max}) até ao próximo *sample* ser disparado, a probabilidade relativa do *sample* que irá ser tocado a seguir e a probabilidade relativa de o processo ser interrompido.

¹⁵ Listas de elementos sonoros passíveis de serem reproduzidos mediante uma determinada sequência.

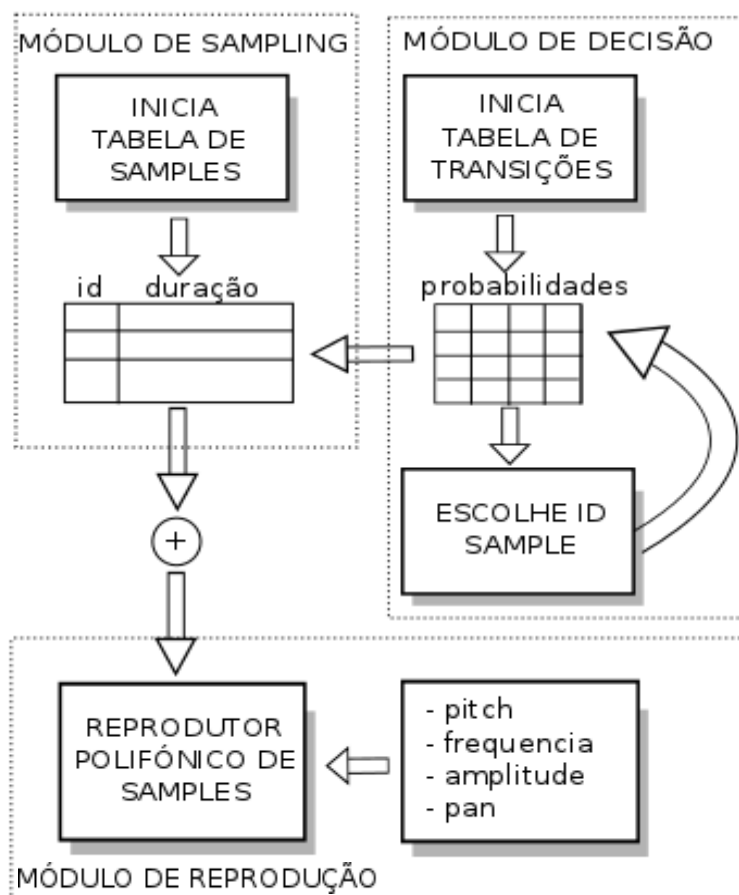


Figura 7. Diagrama do algoritmo

4.2.1 Módulo de sampling

Este módulo organiza uma *playlist* de *samples* que consistem nos conteúdos ordenados que irão ser utilizados no processo generativo das cenas sonoras. Contém um ID que identifica cada *sample* que irá ser utilizado pela tabela de transições e a sua duração que permite ao algoritmo atrasar o *playback* de outro *sample* nessa mesma voz, até o *sample* ter completado o seu ciclo e libertar a voz correspondente.

4.2.2 Módulo de decisão

A geração e organização da sequência de eventos da *playlist* ocorre através de uma tabela de transições que determina a lógica como os eventos se desenvolvem temporalmente, dependendo de relações probabilísticas entre cada um deles. A tabela de transições é utilizada para organizar estas relações probabilísticas garantindo a geração de sequência controladas até certo ponto pelo autor dos cenários sonoros e a não repetição das mesmas.

Atua como um módulo de decisão no sentido em que determina:

- a) **Qual** o *sample* ID que irá ser tocado a seguir baseado no estado actual do *sample* que está a ser tocado. Cada *sample* pode ser considerado como um estado e a tabela contém as probabilidades de cada estado saltar para o estado seguinte.
- b) **Quando** é que a transição de estados acontece através da geração de um número aleatório entre um valor temporal mínimo (*tmin*) e um valor temporal máximo (*tmax*) em unidades de tempo (*msec*).

4.2.3 Módulo de reprodução

Este módulo faz a reprodução dos eventos gerados através da tabela de transições e permite que estes toquem em simultâneo dada a sua implementação polifónica. Este carácter polifónico é necessário de forma a gerir possibilidades em que outro *sample* possa ser iniciado enquanto o *sample* anterior ainda não terminou o *playback*, utilizando uma voz que esteja livre, ou por ação do utilizador, se quiser interferir e reiniciar a rotina.

O algoritmo permite ainda, o controlo manual pelo utilizador de vários parâmetros associados aos eventos, acrescentando, a um dado valor central pré definido, ou definido pelo utilizador, margens de variação (*jitter*) aleatórias.

Estes parâmetros são:

- **Amplitude (Amp)** - controla a amplitude da reprodução do *sample*. O utilizador decide a margem de variação sendo que 0= sem variação e 1= variação máxima:

$$amp = amp_i * amp_{jitter} \text{ onde } amp_{jitter} \in rand(0,1)$$

- **Altura (Pitch)** - Controla a variação em termos de altura de cada *sample* cada vez que é tocado relativamente a um valor inicial que corresponde ao valor por defeito da gravação:

$$pitch = pitch_i * pitch_{jitter} \text{ onde } pitch_{jitter} \in rand(0,1)$$

- **Frequência (Freq)** - Controla a velocidade de reprodução do *sample*. O utilizador decide uma margem de variação em torno do valor original que corresponde ao valor central (valor pré-definido=*sample rate*):

$$Freq = freq_i * freq_{jitter} \text{ onde } freq_{jitter} \in rand(0,1)$$

- **Panorâmica (Pan)** - Controla a espacialização dos *samples* em termos de panorâmica estereofónica. O utilizador decide a posição inicial entre 0 e 1, sendo que 0 corresponde à posição do *sample* totalmente à esquerda e 1 totalmente é direita, e a quantidade de variação relativa a este valor:

$$Pan = pan_i \pm pan_{jitter} \text{ onde } pan_{jitter} \in rand(0,1)$$

- **Reverberação (Rev)** - Controla a relação entre som direto e som reflectido aplicado ao *sample*. O utilizador decide a quantidade de *jitter* aplicado a um valor inicial onde 0= sem variação e 1= variação máxima:

$$Rev = rev_i * rev_{jitter} \text{ onde } rev_{jitter} \in rand(0,1)$$

4.2.4 Exemplo prático

Poderemos imaginar o seguinte cenário sonoro: uma selva amazónica. Este cenário é composto por várias classes de sons que podemos descrever:

- Sons de animais: pássaros, macacos, insectos, etc
- Sons da natureza: vegetação, vento, água, chuva, etc..
- Sons como indicadores: Folhas a mexerem com a atividade dos animais, vento, trovões.

O autor do cenário define as relações probabilísticas de sucessão entre os eventos.

A sequência de eventos poderia ser o som de vários animais localizados em vários pontos da imagem estereofónica, os sons vão surgindo associados a variações de parâmetros que lhes conferem indicadores de localização como referido no capítulo 2 relativo ao espaço do som.

Os eventos relativos aos sons de animais contém alguma probabilidade, enquanto estão a ser reproduzidos, de despoletar um evento diverso como um trovão que funciona como indicador de uma tempestade. Este evento, se acontecer, determina que a probabilidade de os *samples* seguintes serem animais é reduzida mas a probabilidade de o evento seguinte serem *samples* de chuva é elevada e eventualmente existe alguma probabilidade de serem disparados alguns sons bastante espaçados de animais. Estes *samples* de animais, por sua vez, determinam probabilidades nulas de dispararem *samples* de chuva, re combinando e reiniciando o processo de uma forma não repetitiva.

Esta relações probabilísticas contidas na tabela de transições determinam de uma forma narrativa e semi-aleatória, por intermédio de edição do autor, a sequência e relação entre eventos funcionando como um guião sonoro.

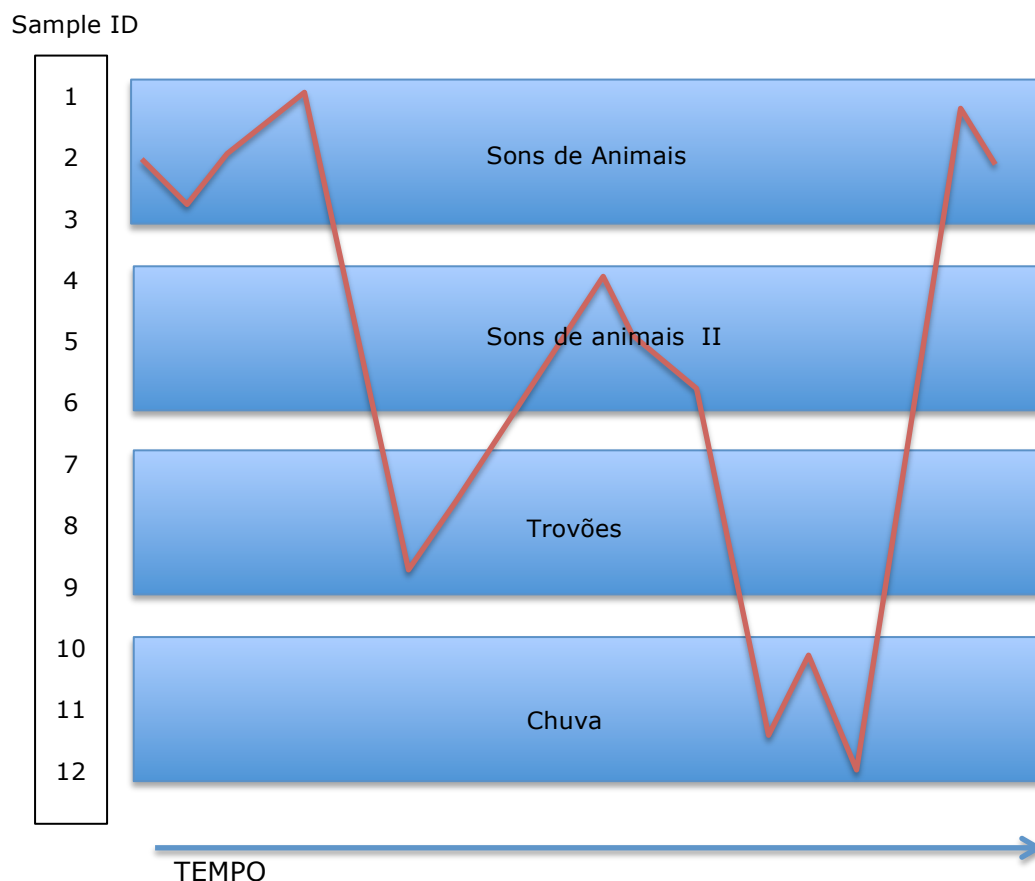


Figura 8. Representação gráfica de uma sucessão de cenas sonoras

4.3 Implementação

A implementação do algoritmo foi feita utilizando as ferramentas MaxMSP e Kyma (Fig. 9). O módulo de *sampling* e o módulo de decisão foram programados no MaxMSP e o módulo de reprodução utilizou a linguagem e o processador Kyma. De uma forma generalista podemos dizer que o MaxMSP controla o Kyma através de uma *network* que utiliza o protocolo *Open Sound Control* (OSC) via *Ethernet*.

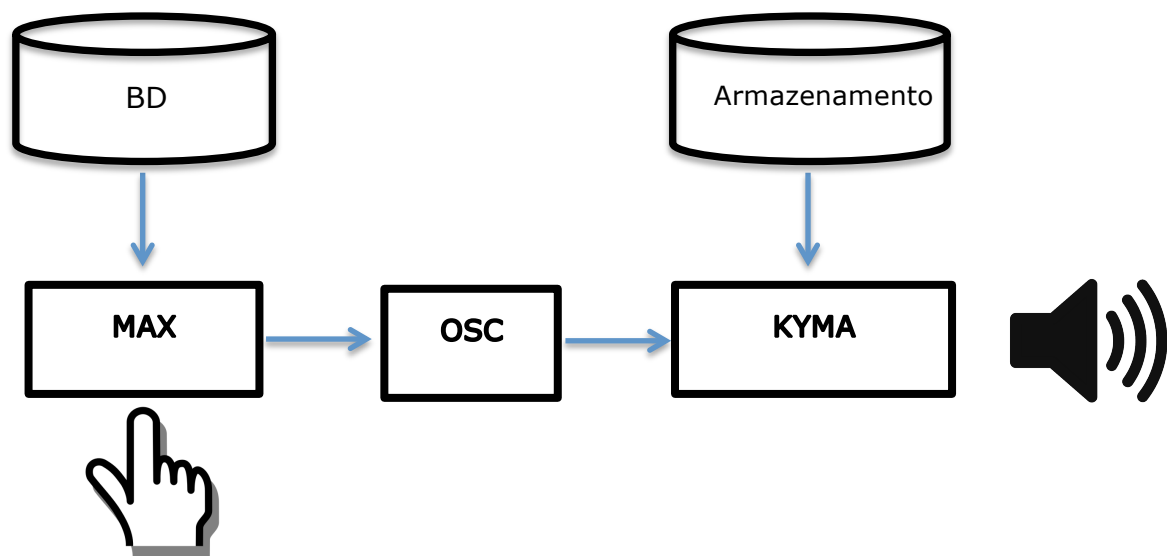


Figure 9. Diagrama de Implementação

Max/MSP - Determina que evento irá ser gerado a seguir e quando irá acontecer através de um método probabilístico que obtém a informação a partir de uma base de dados (Fig.9 “BD”). Os eventos são, neste caso, lds de *samples* sonoros. Os parâmetros de processamento associados a cada evento que são decididos pelo utilizador.

Kyma - faz a reprodução dos *samples* armazenados no disco duro (Fig.9 “Armazenamento”) e o processamento em tempo real do sinal áudio em termos de amplitude, velocidade de reprodução, *pitch*, panorâmica e reverberação.

OSC - estabelece a comunicação entre o *software* transmitindo ao módulo de reprodução os valores gerados pelo módulo de decisão.

4.3.1 Ferramentas

4.3.1.1 Max MSP

“Max is a graphical music programming environment for people who have hit the limits of the usual sequencer and voicing programs for MIDI equipment.” (Puckette, 1988)

O ambiente de programação Max foi criado em 1986 por Miller Puckette no Institute de Recherche et Coordination Acoustique/Musique (IRCAM), em Paris, como ferramenta de criação de música interativa, possibilitando o controlo total da comunicação com outras plataformas suportando vários tipos de protocolos entre os quais MIDI, OSC, etc. [23]

Em 1991 tornou-se um produto comercial, distribuído pela companhia Opcode Systems, com desenvolvimento de Puckette e David Zicarelli, e desde 2000 passou a ser desenvolvido e comercializado pela empresa Cycling '74, em San Francisco criada em 1997 por Zicarelli.

Desde então foram também criadas as expansões ou bibliotecas MSP (*Max Signal Processing*) para processamento de áudio e Jitter para processamento de matrizes, optimizado para a criação e tratamento de vídeo, gráficos 2D e 3D.

O nome do programa é uma homenagem a Max Mathews, investigador nos laboratórios Bell e pioneiro da música por computador.

Baseado na linguagem de programação C, o ambiente Max é uma linguagem *data-flow*, e consiste numa interface gráfica, que permite criar aplicações complexas ligando pequenas “caixas” que desempenham funções específicas.

Grande parte dessas funções são primitivas, semelhantes às aquelas encontradas em todas as linguagens de programação, como operações matemáticas, objetos que implementam estruturas de repetição, estruturas de seleção, etc. E outras menos comuns, adaptadas a toda a lógica e particularidades específicas do Max.

Essas pequenas caixas são combinadas de forma que um programa em Max (*patch* ou *patcher*) é muito semelhante a um fluxograma, tornando mais fácil quer a concepção de um programa, quer o seu *debugging* [24].

4.3.1.2 KYMA

O sistema Kyma é uma combinação de componentes de software e *hardware*. O software é onde a linguagem de programação opera, enquanto o *hardware* serve o primeiro calculando e gerando os sons.

A linguagem de programação Kyma é fundamentalmente uma linguagem visual orientada a objetos, denominados por protótipos e à transmeabilidade de informação, fornecendo diversos caminhos para a visualizar, obter e manipular com o objectivo de forjar caminhos sonoros e musicais. Podemos considerar esta linguagem vinculada a uma domínio específico, neste caso o musical, no sentido em que foi desenhada para facilitar a síntese e a transformação sonora através da exploração e construção de sons e performance em tempo real. É uma linguagem baseada em *Smalltalk* com

uma variação adaptada à computação em tempo real e dinâmica denominada por *Capytalk* [25]

O *software* Kyma não denomina por *patch* as suas estruturas mas sim por sons. Deste modo, todos os pequenos elementos denominados por protótipos (fig. 10) são considerados sons e o produto da sua interligação é considerado um som final ainda que porventura mais complexo e extenso.



Figura 10. Exemplo de um protótipo do Kyma

A aceleração áudio à executada através do *hardware* que o sistema utiliza denomina por Pacarana (fig.11) e tem como função libertar o computador de processamento paralelo como o gráfico ou de sistema operativo, o que permite dedicar cada ciclo de processamento ao som.

A utilização destes processadores dedicados permite ao sistema, produzir em tempo real milhares de osciladores em simultâneo e dezenas de efeitos como reverberação ou convolução.



Figura 11. Pacarana: o sistema de DSP mais recente da *symbolic sound*.

4.3.1.3 OpenSound Control

O OSC (*OpenSound Control*) é um protocolo desenvolvido pelo CNMAT (*Center for New Music and Audio Technology*) da Universidade de Berkeley, Califórnia e, à semelhança do protocolo MIDI, tem como propósito possibilitar a comunicação entre computadores, sintetizadores, e outros dispositivos multimédia. A especificação deste protocolo - *OpenSound Control Specification 1.0* - foi redigida em 2002 e, sendo cer-

ca de 20 anos mais recente que o MIDI, revela-se mais apropriado às necessidades atuais, nomeadamente, na utilização de redes de comunicação modernas para a transferência de dados, e na definição de uma linguagem com maior abertura e maior resolução do que a utilizada no protocolo MIDI. [26]

Resumo das características do OSC

A arquitetura do OSC é baseada no conceito de *cliente / servidor*, sendo que qualquer dispositivo que envie pacotes é considerado cliente e qualquer outro que os receba, será considerado servidor. Um dispositivo poderá ser considerado cliente e servidor simultaneamente. O OSC é um protocolo completamente independente da tecnologia física de transporte da informação. Desta forma, torna-se extremamente versátil, podendo-se integrar em praticamente qualquer tecnologia de rede.

No OSC, o método de atribuição de nomes às mensagens é feito de forma dinâmica e aberta, com uma estrutura semelhante à de um URL (ex: “index1/4/freq”). São suportados vários tipos de argumentos para as mensagens: *strings* ASCII, *ints* e *floats* de 32 bit, 64 bit, entre outros. Este protocolo incorpora um sistema de pergunta (queries) para conhecer as capacidades de outro dispositivo OSC de uma forma dinâmica. Possibilita ainda a utilização de funções de *Pattern Matching* (“?”, “*”, “[]”, etc.) com utilização similar à de uma *shell* em UNIX, permitindo assim a especificação de múltiplos destinos para uma única mensagem[24]

Em termos de gestão temporal, a inclusão de *Time Tags* permite uma alta precisão temporal (abaixo do nano- segundo), utilizando o mesmo formato do protocolo Internet *Network Time Protocol* (NTP). Contempla ainda a possibilidade do envio de *Bundles* de mensagens (com um *time tag* associado), cuja ação deva ocorrer simultaneamente.

4.3.1.4 Processamento distribuído

Apesar das crescentes capacidades dos computadores ao nível do processamento, seja através do aumento da frequência de relógio, da adição de processadores (ex: *Dual* ou *Quad Processing*) ou da própria arquitetura dos mesmos (ex: *Dual Core*, 64-bit), a realidade é que também nós, enquanto utilizadores dessa mesma tecnologia, exigimos mais dos equipamentos, passando a utilizar os mesmos em contextos até então não considerados. Daí que, por muito capaz que seja o poder de processamento dos computadores, por vezes tenhamos que recorrer a técnicas de processamento distribuído. O sistema proposto nesta tese, devido à sua potencial complexidade, deverá ser o mais flexível possível, contemplando a necessidade de distribuir o processamento por vários computadores, neste caso dois: o ambiente Max MSP e o ambi-

ente Kyma, sendo que, no caso do ultimo, o hardware que lhe é intrínseco, é por si só, um computador dedicado.

A comunicação entre os vários elementos da rede informática é assegurada através de qualquer tecnologia de rede que suporte a arquitetura protocolar TCP/IP (redes locais sem fios, *Ethernet*, *Firewire*, etc). Uma vez que esta arquitetura é atualmente a mais implementada em termos de indústria, este facto confere uma grande abertura ao sistema proposto, permitindo a este funcionar até através da própria *Internet*. Através da tecnologia de rede utilizada poderão ser transmitidas informações OSC, MIDI ou mesmo áudio digital. No decorrer da implementação do sistema proposto, utilizámos a comum tecnologia *Ethernet* a 100 Mbps para a interligação de computadores. Uma vez que as necessidades de transferência do sistema em termos de débito binário não são consideráveis, esta revelou-se extremamente fiável. É conveniente referir ainda que, funcionando o sistema em tempo-real, ao escolher a tecnologia de rede a ser implementada será fundamental considerar também o atraso (latência) e a variação do atraso (*jitter*) típicos da mesma, parâmetros tão importantes neste contexto como o próprio débito binário [27].

4.3.2 Desenvolvimento da aplicação

4.3.2.1 Max/MSP - *Patch* principal

O *patch* principal da aplicação (fig. 12), “alberga” todos os elementos que compõem o interface do utilizador e todos os *subpatches* que contém as rotinas necessários ao funcionamento do algoritmo: módulo de decisão (decide o *sample* que vai ser tocado a seguir), parâmetros de reprodução e comunicação através da *network* com o *software Kyma*.

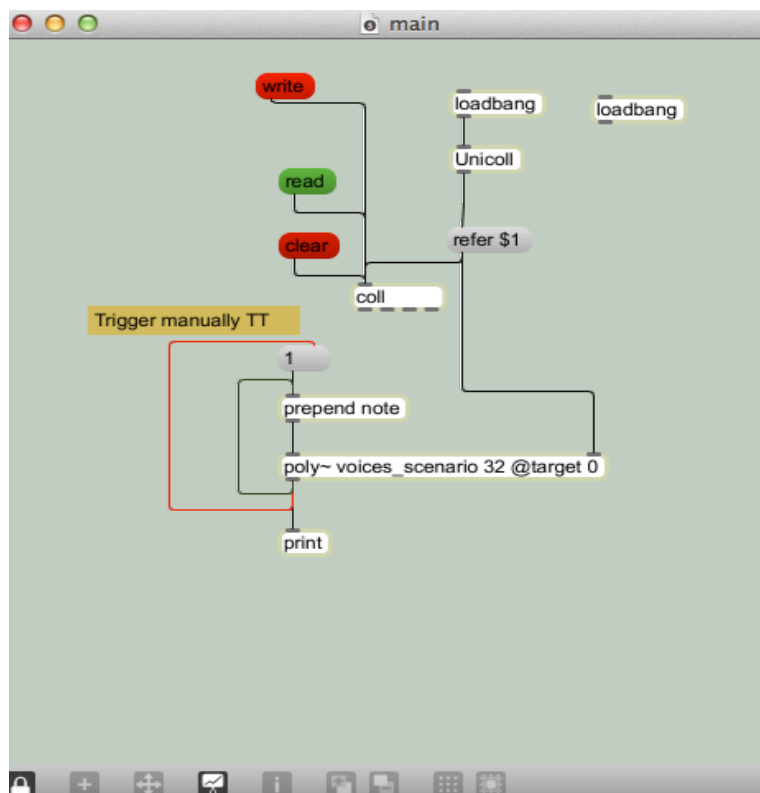
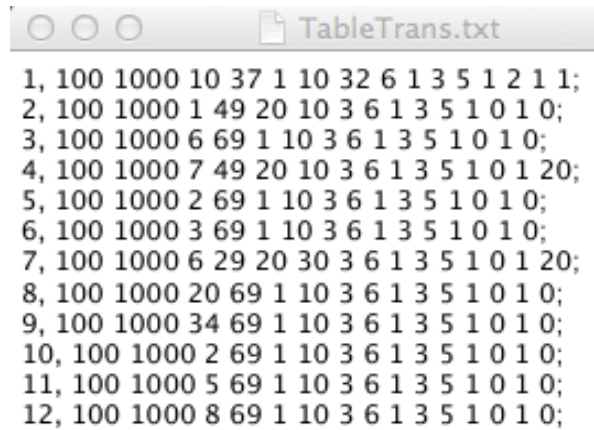


Figura 12. Patch principal da aplicação

O utilizador inicia manualmente neste *patch* o processo generativo e pode atribuir valores aos parâmetros de *jitter* descritos no capítulo 4.3.2 referente à descrição do algoritmo. É ainda no interior deste *patch* que tem a opção de carregar manualmente a tabela de transições pretendida e que está armazenada num ficheiro de texto no disco duro. A opção por esta metodologia prendeu-se com motivos de coerência perante a concepção do projeto, uma vez que se pretende que o autor dos cenários sonoros construa um guião de eventos que pode alterar a qualquer momento. Deste modo e idealmente, após ter o domínio da ferramenta, pode ir compilando vários cenários à medida que vai experimentando novas combinações e relações probabilísticas entre os sons que pode utilizar perante objectivos diversos.

Este ficheiro de texto (fig.13) é armazenado no interior de um objecto *coll* e a sua estrutura está organizada da seguinte forma:

A colunas 1 e 4 estão relacionadas com o módulo de *sampling* descrito no capítulo referente à estrutura do algoritmo (*Sample ID* e duração do *sample* em milissegundos). As colunas 2 e 3 estão relacionadas com um intervalo de tempo (*tmin* e *tmax*) que delimita a janela temporal entre a qual ocorrerá um novo evento. Finalmente as restantes colunas, 5 até à ultima coluna (dependendo do número de *samples* que compõe a matriz), determinam as probabilidades de cada evento ocorrer, após a reprodução do evento atual e a probabilidade do processo terminar.



```

1, 100 1000 10 37 1 10 32 6 1 3 5 1 2 1 1;
2, 100 1000 1 49 20 10 3 6 1 3 5 1 0 1 0;
3, 100 1000 6 69 1 10 3 6 1 3 5 1 0 1 0;
4, 100 1000 7 49 20 10 3 6 1 3 5 1 0 1 20;
5, 100 1000 2 69 1 10 3 6 1 3 5 1 0 1 0;
6, 100 1000 3 69 1 10 3 6 1 3 5 1 0 1 0;
7, 100 1000 6 29 20 30 3 6 1 3 5 1 0 1 20;
8, 100 1000 20 69 1 10 3 6 1 3 5 1 0 1 0;
9, 100 1000 34 69 1 10 3 6 1 3 5 1 0 1 0;
10, 100 1000 2 69 1 10 3 6 1 3 5 1 0 1 0;
11, 100 1000 5 69 1 10 3 6 1 3 5 1 0 1 0;
12, 100 1000 8 69 1 10 3 6 1 3 5 1 0 1 0;

```

Figura 13. estrutura do ficheiro de texto

4.3.2.2 Max/MSP - Abstraction “Cenário”

O *patch* “cenário” (fig. 14) relaciona os 3 módulos do algoritmo na medida em que recebe e armazena os dados da tabela (módulo de *sampling*), decide qual o *sample* que irá ser disparado (módulo de decisão) e organiza a informação a ser enviada por OSC para reprodução (módulo de reprodução).

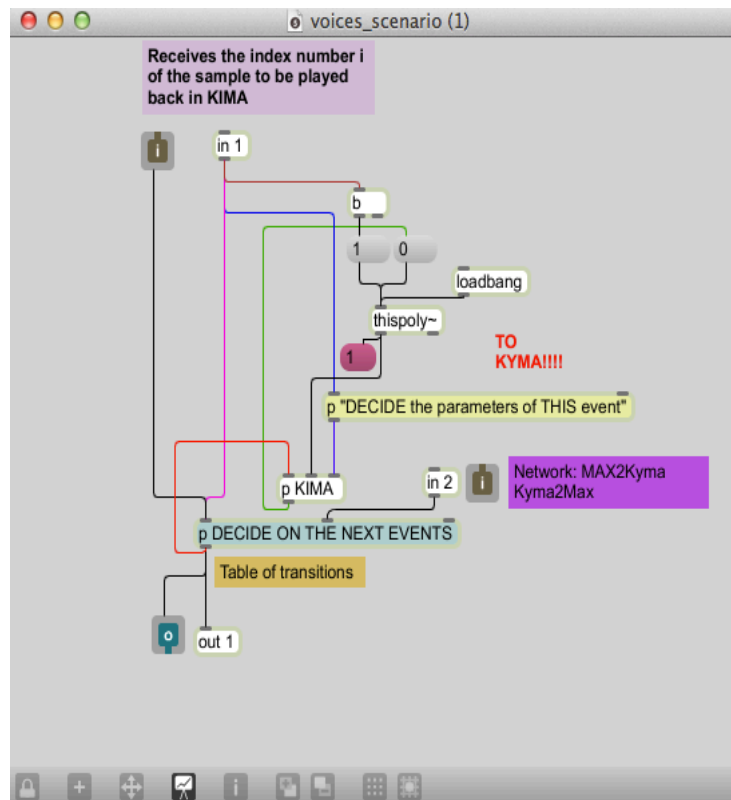


Figura 14. Max/MSP Abstraction “Cenário”

de informação gerado no *subpatch* “cenário” de forma a enviar os dados para o *software* Kyma através de mensagens OSC (fig. 16).



Figura 16. Tipologia de uma mensagem OSC: prefixo, parâmetro, ID e valor.

As mensagens OSC enviadas são do tipo *strings* ASCII e estruturam-se do seguinte modo: [Prefixo]/[parâmetro]/[ID] [valor]

- **Prefixo** - no caso de pretendemos utilizar mais do que 1 cenário em simultâneo, necessitamos de atribuir um ID único às mensagens de forma a serem discretizadas e mapeadas independentemente no Kyma.
- **Parâmetro** - Atribui um nome ao parâmetro que será alocado pelo Kyma ao controlador correspondente.
- **ID do *sample*** - associa os parâmetros aos *samples* correspondentes.
- **Valor** - Refere-se ao valor do parâmetro para o *sample* que está a ser reproduzido obtido através da introdução do valor de *jitter* pelo utilizador.

4.3.2.3 Kyma - Som principal

Como foi descrito no subcapítulo 4.3 dedicado à implementação do algoritmo, o sistema Kyma foi utilizado para a reprodução e processamento áudio dos *samples* sonoros após a sequência e parâmetros decididos no Max/MSP.

As razões para esta distribuição de funções, prenderam-se com as capacidades que o processamento dedicado proporciona em termos de qualidade do áudio, a expansão da aplicação que pretendemos para o futuro e o facto de existirem dentro deste sistema várias soluções pré-concebidas que facilitam o processo de elaboração de sons com óptima qualidade.

A implementação no aplicação Kyma consistiu numa primeira fase, em construir uma espécie de espelho de todos os eventos decididos pelo Max/MSP: *sample* a ser tocado a seguir, amplitude, frequência, *pitch*, panorâmica e reverberação de forma a que as vozes pudessem ser processadas individualmente (Fig.17).

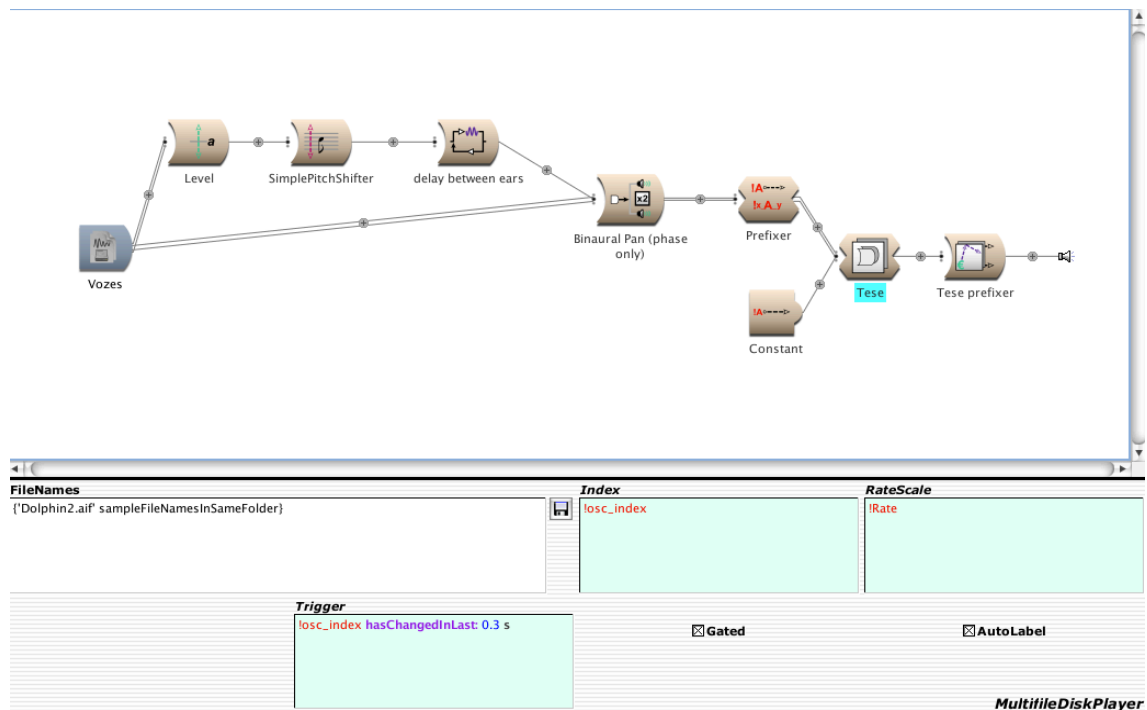


Figura 17. Fluxo de áudio e ligação de protótipos no sistema Kyma

Desta maneira os principais aspectos a considerar foram a relação entre os valores gerados no MAX/MSP em termos de *samples* a serem reproduzidos e parâmetros associados e o sincronismo na resposta dos sons gerados no Kyma. À luz disto, houve a necessidade de fazer comunicar cada um dos controladores dos protótipos do Kyma com cada uma das vozes iniciadas no Max/MSP através da atribuição de *IDs* únicos na estrutura das mensagens OSC coincidente com a atribuição de um prefixo numérico a cada um dos parâmetros do Kyma.(Fig.18)



Figura 18. Parâmetro de controlo de um protótipo do Kyma através de OSC

Uma vez que o objecto *poly~* do Max desempenha funções polifónicas houve a necessidade de criar a mesma possibilidade no Kyma através de um protótipo denominado por *Replicator* que cria múltiplas instâncias da mesma voz permitindo o controlo de cada uma das vozes individualmente.

4.3.2.4 Implementação do VCS (Virtual Control Surface)

A fase de construção de sons no ambiente Kyma é distinta do ambiente de controlo dos mesmos uma vez que o sistema faz a compilação de grande parte dos algoritmos anteriormente à sua possibilidade de manipulação em tempo real. Ou seja, após construirmos o esquema de ligações e fluxo de sinal entre os protótipos, o sistema compila a informação que surge graficamente na forma do VCS (Fig.17) o que permite um sistema híbrido de monitorização de sinal e manipulação do mesmo em tempo real. É também nesta estrutura que podemos organizar a superfície em termos de interface para facilitar a interatividades com o utilizador organizando o espaço gráfico, as dimensões dos controladores, decidindo acerca do que é considerado importante em termos de visibilidade, etc.

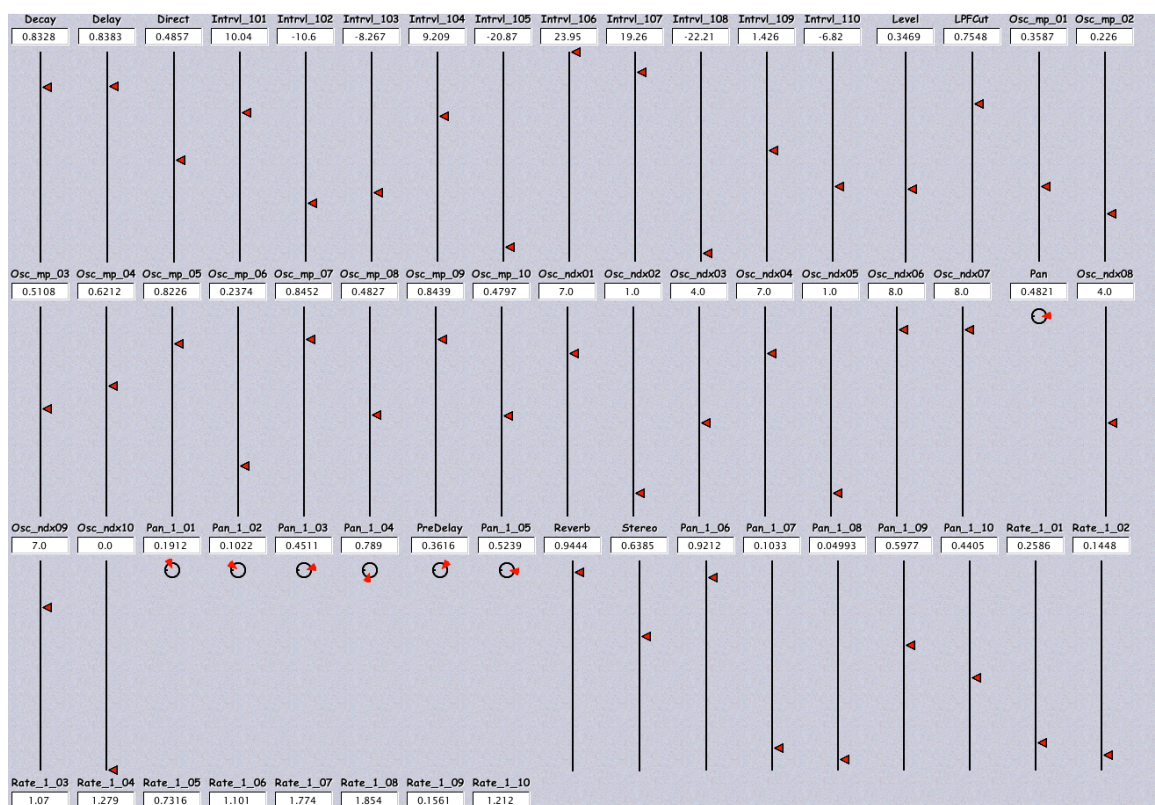


Figura 19. VCS (Virtual Control Surface)

Capítulo 5

5. Conclusões e Perspectivas de Desenvolvimento

No contexto do tema subjacente a esta tese, sabíamos de antemão a dificuldade que iríamos encontrar em validar, pelo menos nesta fase do trabalho os resultados obtidos, não em termos técnicos ou de implementação, mas em termos de eficácia obtida aferida pelo julgamento de uma amostra significativa de terceiros. Estas dificuldades foram-se tornando evidentes desde o primeiro capítulo quando a discussão teórica indica, de um modo geral, o carácter multimodal e híbrido do conceito de imersão sonora.

Se tivemos a pretensão de sugerir a inclusão de modelos mais subjetivos que também constituem, a nosso ver o processo imersivo, não tivemos a pretensão de estabelecer as correlações ideais entre o estado imersivo e o domínio sonoro. Tivemos, no entanto, o objectivo de alargar o âmbito de análise e sugerir um protótipo salientando o potencial tecnológico de algumas ferramentas sonoras e capaz de incluir hipóteses discutidas durante a discussão teórica. Sustentamos que a manipulação de certos parâmetros sonoros poderá encontrar uma correspondência cognitiva e interpretativa e que o seu controlo poderá potenciar níveis de imersão áudio.

Em nosso entender, a investigação acerca da imersão sonora encontra-se dividida em dois grupos: por um lado a vertente de engenharia de áudio e estudos psicoacústicos acerca de técnicas de espacialização áudio e por outro o âmbito da emulação realista de ambientes sonoros com recurso a tecnologia de computação e de síntese sonora.

Pensámos que o campo de investigação ligado à psicologia do ouvinte e à manipulação criativa do som é um campo teórico a ser explorado e que poderá funcionar como um eixo entre os dois grupos mencionados.

Após análise ao percurso traçado no âmbito desta tese, somos da opinião que o conhecimento gerado é apenas um ponto de partida uma vez que as possibilidades de extensão deste tipo de abordagens são bastante abrangentes e existem várias arestas

a limar. Gostaríamos de aproveitar este capítulo para expor algumas destas avaliações.

Avaliação crítica

O protótipo do sistema sugerido, apesar de funcional, é rígido, no sentido em que foi implementado com um objectivo muito particular: o de criar uma aplicação capaz de gerar automaticamente sequências sonoras e construir cenários sonoros à luz de uma perspectiva de autor. Esta perspectiva consistiu em operacionalizar certas hipóteses teóricas, as quais, após terem sido analisadas, inspiraram o projeto.

O autor conceptualiza o cenário, dita as regras em termos de seleção de sons e relação probabilística entre os mesmos e inicia um processo automático e sequencial aproveitando o potencial tecnológico de gestão e transformação da informação. Neste sentido vemos como principais contribuições do projeto:

- A possibilidades da sua utilização em áreas diversas como os jogos, cinema, instalações multimédia, simuladores (arquitectónicos, de voo, acústica ambiental, turismo, jogos sérios, etc.). A título de exemplo gostaríamos de acrescentar que a EA Sports e a Sony vão integrar nos seus motores de áudio futuros aplicações do tipo Max/MSP como o Pure Data de fora a tirarem partido do crescente interesse e potencial de técnicas de síntese e de áudio processual.
- Pode ser utilizado de uma perspectiva mais ou menos realista consoante o objectivo do autor e a complexidade que pretende implementar nos cenários sonoros e mais ou menos aleatória, também consoante o tipo de cenário pretendido.
- Possibilita a criação de soluções personalizadas devido ao controlo que o autor tem sobre os cenários sonoros nomeadamente na conceptualização de cenas e seleção de sons.
- Simplifica o trabalho de *authoring*.

O protótipo pretendido, como foi referido anteriormente, teve como principal motivação, a operacionalização de alguns conceitos obtidos através da investigação teórica, pelo que os objectivos em realizar uma aplicação distribuída, autónoma e flexível em termos de meios tecnológicos, não estiveram, ainda, no topo dos objectivos. À luz disto, poderemos dizer que uma das pretensões futuras será estender esta investigação e implementação a outro nível que contemple mais adaptabilidade e funcionalidade, sendo que desenvolver a aplicação dentro de um só software ou sistema será desejável. Outro dos pontos que gostaríamos de otimizar será a simplificação na autoria dos cenários uma vez que o processo manual que implica a determinação dos valores probabilísticos entre eventos é ainda muito moroso e trabalhoso.

Gostaríamos de concluir salientando que, apesar de ser um campo de estudo relativamente novo, o simples facto de explorar a relação entre áudio e o estado de imersão oferece inúmeras oportunidades exploratórias e criativas que julgamos terem bastante sentido, tanto do ponto de vista estético e artístico, como do ponto de vista da investigação científica. Como muitos autores referiram anteriormente nas ideias introdutórias desta tese, o carácter subjetivo e de difícil análise do tema, pelo menos em termos do som, pode torná-lo algo opaco no sentido de ser difícil quantificar certos parâmetro que se julga fazerem parte da equação complexa que é um ambiente imersivo. No entanto, no contexto dos novos media e da sua vertente mais criativa, a relevância do tema e as questões que coloca, julgámos serem de todo o interesse para o âmbito de investigação desta comunidade.

6. Referências Bibliográficas

- [1] Ermi, L. & Mäyrä, F. (2005). *Fundamental Components of the Gameplay Experience: Analysing Immersion* [versão electrónica].
- [2] Nachvatal, Joseph (2011). "Immersion Into Noise", MPublishing, University of Michigan Library
- [3] [http://en.wikipedia.org/wiki/Immersion_\(virtual_reality\)](http://en.wikipedia.org/wiki/Immersion_(virtual_reality)). Acesso a 11 de Abril de 2011
- [4] Dovey, J. And Kennedy, H.W. (2006) *Game Cultures: Computer games as new Media*. Berkshire: Open University Press.
- [5] Carr, D. (2006). "Space, Navigation and Affect." In *Computer Games: Text, Narrative and Play*, Cambridge: Polity.
- [6] Taylor, L. N. (2003). *When seams fall apart: Video game space and the player*. Game Studies, 3(2). <http://gamestudies.org/0302/taylor/>. Acesso a 12 de Abril de 2011
- [7] Björk, Staffan; Jussi Holopainen. (2004). *Patterns In Game Design*. Charles River Media. (versão digital)
- [8] Calleja, G. (2007b). *Digital Games as Designed Experience: Reframing the Concept of Immersion*. Victoria University of Wellington. (versão digital)
- [9] Brown, E. & Cairns, P. (2004). *A Grounded Investigation of Game Immersion*. University College London Interaction Centre (UCLIC). Conference on Human Factors in Computing Systems. Vienna, Austria. (versao digital)
- [10] [http://en.wikipedia.org/wiki/Immersion_\(virtual_reality\)](http://en.wikipedia.org/wiki/Immersion_(virtual_reality))
- [11] Durand, R. Begault. "3-D Sound for Virtual Reality and Multimedia", NASA Center for AeroSpace Information (Abril 2000)
- [12] Murray Schafer, R. (1977, 1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, Vermont: Destiny Book.
- [13] Andrea Valle, Vincenzo Lombardo, and Mattia Schirosa. *A graph-based system for the dynamic generation of soundscapes*. In Proceedings of the 15th International Conference on Auditory Display, pages 217-224, Copenhagen, 2009. (versão digital)

- [14] Chion, M. (1994). *Audio-Vision: Sound on Screen*. C. Gorbman, trans. New York: Columbia University Press.
- [15] Grimshaw, Mark. (2007). *The resonating spaces of First-Person Shooter games. Games Computing and Creative Technologies: Conference Papers* Paper 4. (versão digital)
- [16] LaBelle, B. & Martinho, C. (2011). *Site of Sound: Of Architecture and the Ear, Volume 2*, Errant Bodies Press
- [17] Henrique, L. (2002). *Acústica Musical*: Fundação Calouste Gulbenkian.
- [18] Schwartz, E and BG Shinn-Cunningham (2010). "Dissociation of perception judgments of 'what' and 'where' in ambiguous auditory scenes," Journal of the Acoustical Society of America, in press.
- [19] <http://pt.wikipedia.org/wiki/Cognição>. Acesso a 23 de Maio de 2012
- [20] Lastra, J. Reading (1985). *Writing, and Representing Sound*, In Altman, R. (ed.) Sound theory, Sound practise, Routledge
- [21] Ekman, I. (2009). "Modelling the Emotional Listener: Making Psychological Processes Audible", Proc. AudioMostly 2009, September, Glasgow, Great Britain.
- [22] Leeuwen, T. van. (1999). *Speech, Music, Sound*. Macpress, London.
- [23] <http://cycling74.com/whatismax/>. Acesso a 27 de Junho de 2012
- [24] Dias, R. (2009). *Plataforma modular para prototipagem de sistemas interativos*. Dissertação de Mestrado, FEUP
- [25] Stolet, J. (2011). *Kyma and the SumOfSines Disco Club*. Jeffrey Stolet (Standard Copyright License)
- [26] Santos, P. (2009). *Tecnologias de visualização sonora num contexto artístico de performance musical*. Dissertação de Mestrado, FEUP
- [27] http://en.wikipedia.org/wiki/Sound_design. Acesso a 25 de Maio de 2012
- [28] http://www.edtl.com.pt/index.php?option=com_mtree&task=viewlink&link_id=1527&Itemid=2. Acesso a 24 de Maio de 2012
- [29] Nogueira, L. (2010). *Planificação e Montagem - Manuais de Cinema 3*. Livros LabCom
- [30] http://en.wikipedia.org/wiki/Concatenative_synthesis. Acesso a 24 de Maio de 2012
- [31] <http://www.sonicspot.com/guide/synthesistypes.html>. Acesso a 24 de Maio de 2012

7. Anexo A

R. Murray Schafer's Taxonomy

1. Natural Sounds

- Sounds of creation - Sounds of apocalypse - Sounds of water - Sounds of air - Sounds of earth - Sounds of fire - Sounds of birds - Sounds of animals - Sounds of insects - Sounds of fish and sea creatures - Sounds of seasons

2. Human sounds

- Sounds of the voice - Sounds of the body - Sounds of clothing

3. Sounds and society

- General description of rural soundscape - Town soundscapes - City soundscapes - Maritime soundscapes

- Domestic soundscapes - Sounds of trades, professions and livelihoods - Sounds of factories and oces - Sounds of entertainment - Music - Ceremonies and festivals - Parks and gardens - Religious festivals

4. Mechanical Sounds

- Machines - Industrial and factory equipment - Transportation machines - Warfare machines - Trains and trolleys - Internal combustion engines- Aircraft - Construction and demolition equipment - Mechanical tools - Ventilations and air-conditioners - Instruments of war and destruction - Farm machinery

5. Quiet and Silence

6. Sounds as Indicators

Bells and gongs - Horns and whistles - Sounds of time - Telephones - Warning systems - Signals of pleasure - Indicators of future occurrence